

# THE SINE-GORDON EQUATION IN THE SEMICLASSICAL LIMIT: CRITICAL BEHAVIOR NEAR A SEPARATRIX

ROBERT J. BUCKINGHAM AND PETER D. MILLER

**ABSTRACT.** We study the Cauchy problem for the sine-Gordon equation in the semiclassical limit with pure-impulse initial data of sufficient strength to generate both high-frequency rotational motion near the peak of the impulse profile and also high-frequency librational motion in the tails. Subject to suitable conditions of a general nature, we analyze the fluxon condensate solution approximating the given initial data for small time near points where the initial data crosses the separatrix of the phase portrait of the simple pendulum. We show that the solution is locally constructed as a universal curvilinear grid of superluminal kinks and grazing collisions thereof, with the grid curves being determined from rational solutions of the Painlevé-II system.

## CONTENTS

1. Introduction	2
1.1. Notation and terminology.	3
1.2. Assumptions and definition of fluxon condensates.	4
1.3. Statement of results.	10
2. Choice of $g(w)$	15
3. Steepest Descent, the Outer Model Problem, and its Solution	18
4. Inner Model Problem Valid near $w = -1$	20
4.1. Exact jump matrices for $\mathbf{P}(w)$ near $w = -1$ .	20
4.2. Expansion of $k(w)$ about $w = -1$ .	21
4.3. Conformal mapping near $w = -1$ and new spacetime coordinates.	24
4.4. Formulation of an inner model problem.	28
5. Solution of the Inner Model Problem	30
5.1. Symmetry analysis of the inner model problem.	30
5.2. Differential equations derived from the inner model problem. Lax equations and Painlevé-II	31
5.3. Solution of Riemann-Hilbert Problem 4.1 for $m = 0$ .	34
5.4. Schlesinger-Bäcklund transformations. Solution of Riemann-Hilbert Problem 4.1 for general $m \in \mathbb{Z}$ .	36
6. Inadequacy of the Global Model	38
6.1. Parametrix for $\mathbf{Q}_m(w)$ in $\Omega_m^+$ .	42
6.2. Parametrix for $\mathbf{Q}_m(w)$ in $\Omega_m^-$ .	47
6.3. Accuracy of the parametrices for $\mathbf{Q}_m(w)$ . Error analysis.	49
7. Proofs of Theorems	51
7.1. Exact formulae for the potentials.	51
7.2. Proof of Theorem 1.1 (Initial accuracy of fluxon condensates).	54
7.3. Proof of Theorem 1.2 (Main approximation theorem).	55
7.4. Proof of Theorem 1.3 (Superluminal kink asymptotics).	57
7.5. Proof of Theorem 1.4 (Grazing kink collisional asymptotics).	59
References	60

*Date:* June 29, 2011

The authors thank A. B. J. Kuijlaars and A. R. Its for useful discussions. R. J. Buckingham was partially supported by the Charles Phelps Taft Research Foundation. P. D. Miller was partially supported by the National Science Foundation under grant DMS-0807653.

## 1. INTRODUCTION

This paper is concerned with a detailed local analysis of the solution of the Cauchy initial-value problem for the sine-Gordon equation

$$\epsilon^2 u_{tt} - \epsilon^2 u_{xx} + \sin(u) = 0, \quad u(x, 0) = F(x), \quad \epsilon u_t(x, 0) = G(x), \quad x \in \mathbb{R}. \quad (1.1)$$

We will consider the number  $\epsilon > 0$  to be a small parameter. This type of scaling can be physically motivated in the situation that the sine-Gordon equation is used to model the propagation of magnetic flux along superconducting Josephson junctions [23]. The sine-Gordon equation can also be derived in the continuum limit as a model for an array of coaxial pendula with nearest-neighbor torsion coupling [3]. This latter application is particularly useful for the purposes of visualization of solutions.

A dramatic separation of scales occurs in the *semiclassical limit*  $\epsilon \downarrow 0$  if the initial data  $F(\cdot)$  and  $G(\cdot)$  are held fixed. As can be seen in Figures 1.1–1.3, the semiclassical dynamics apparently consists of well-defined (asymptotically independent of  $\epsilon$ ) spacetime regions containing oscillations on space and time scales proportional to  $\epsilon$  but modulated over longer scales originating with the  $\epsilon$ -independent initial conditions. An important role is played by the  $x$ -parametrized curve  $(F, G) = (F(x), G(x))$  in the phase portrait of the simple pendulum, and its relation to the separatrix curve  $(1 - \cos(F)) + \frac{1}{2}G^2 = 2$ . Indeed, in the specific context of suitable initial data of *pure impulse* type, that is, for which  $F(x) \equiv 0$ , the following dichotomy has recently been established [7] regarding the asymptotic behavior of the solution  $u(x, t)$  of (1.1) for small time  $t$  independent of  $\epsilon$ . If  $(F, G) = (0, G(x))$  lies inside the separatrix, then  $u(x, t)$  is accurately modeled by a modulated train of superluminal librational waves; but if  $(F, G) = (0, G(x))$  lies outside the separatrix, then  $u(x, t)$  is instead accurately modeled by a modulated train of superluminal rotational waves. If  $x \in \mathbb{R}$  is a point lying exactly on the separatrix curve, then the approximation theorems proved in [7] fail to provide a uniform description of the asymptotic behavior of the solution  $u(x, t)$  near such  $x$ . It is clear from plots of exact solutions shown in Figures 1.1–1.3 that some essentially different asymptotic behavior is generated by separatrix crossings in the pure-impulse initial data. In particular, a different and more complicated kind of waveform than modulated traveling waves appears to spread in time away from specific points where separatrix crossings occur in the initial data. The coupled pendulum interpretation is useful here: if the pendula are all initially at rest in the gravitationally stable configuration and are given a spatially-localized initial impulse of sufficient strength, then some pendula have sufficient energy to rotate completely about the axis a number of times, while the pendula in the “wings” experience very little initial impulse and only have energy for small oscillatory motions near equilibrium (so-called librational motion). Clearly this situation leads to kink generation near the transition points and a more complicated type of dynamics as the kinks struggle to separate from one another.

To better understand the reason for the locally complicated behavior, it is useful to view the sine-Gordon equation as a perturbation of the simple pendulum ODE:

$$u_{TT} + \sin(u) = \epsilon^2 u_{xx}, \quad u(x, 0) = 0, \quad u_T(x, 0) = G(x), \quad t = \epsilon T. \quad (1.2)$$

We might expect that the right-hand side should be negligible for quite a long rescaled time  $T$  as long as in the unperturbed system ( $\epsilon = 0$ ) pendula located at nearby values of  $x$  follow nearby orbits of the pendulum phase portrait. This will be the case unless  $|G(x)| \approx 2$ , the condition allowing for nearby values of  $x$  to correspond to topologically dissimilar orbits, leading to large relative displacements in  $u$  over finite  $T$  and causing  $u_{xx}$  to become very large very quickly. See Figure 1.4. To leading order, and for times  $t$  of order  $\epsilon$  we expect that orbits near  $x = x_{\text{crit}}$  should follow the pendulum separatrix:  $\cos(u/2) \approx \text{sech}(t/\epsilon)$ ,  $\sin(u/2) \approx -\tanh(t/\epsilon)$ , and  $\epsilon u_t \approx -2 \text{sech}(t/\epsilon)$ . Of course as  $T = t/\epsilon \rightarrow \infty$ , the perturbation term  $\epsilon^2 u_{xx}$  will become important and will prevent the rapid divergence of trajectories shown in Figure 1.4.

The subject of this paper is the asymptotic analysis, in the semiclassical limit  $\epsilon \downarrow 0$ , of the solution of the Cauchy problem (1.1) for the sine-Gordon equation in the case that  $t$  is small and  $x$  is near a value where the curve  $(F, G) = (0, G(x))$  crosses the separatrix (at its midpoint, as in Figure 1.4). This region of the  $(x, t)$ -plane is blown up with what turn out to be the correct scalings for easier viewing in Figure 1.5. Therefore we are near the boundary of the regions where the solution is described in terms of modulated rotational and librational solutions of the pendulum equation. We will show that, under suitable further assumptions, the asymptotic behavior of  $u(x, t)$  in this situation is universal, and is described by an essentially multiscale

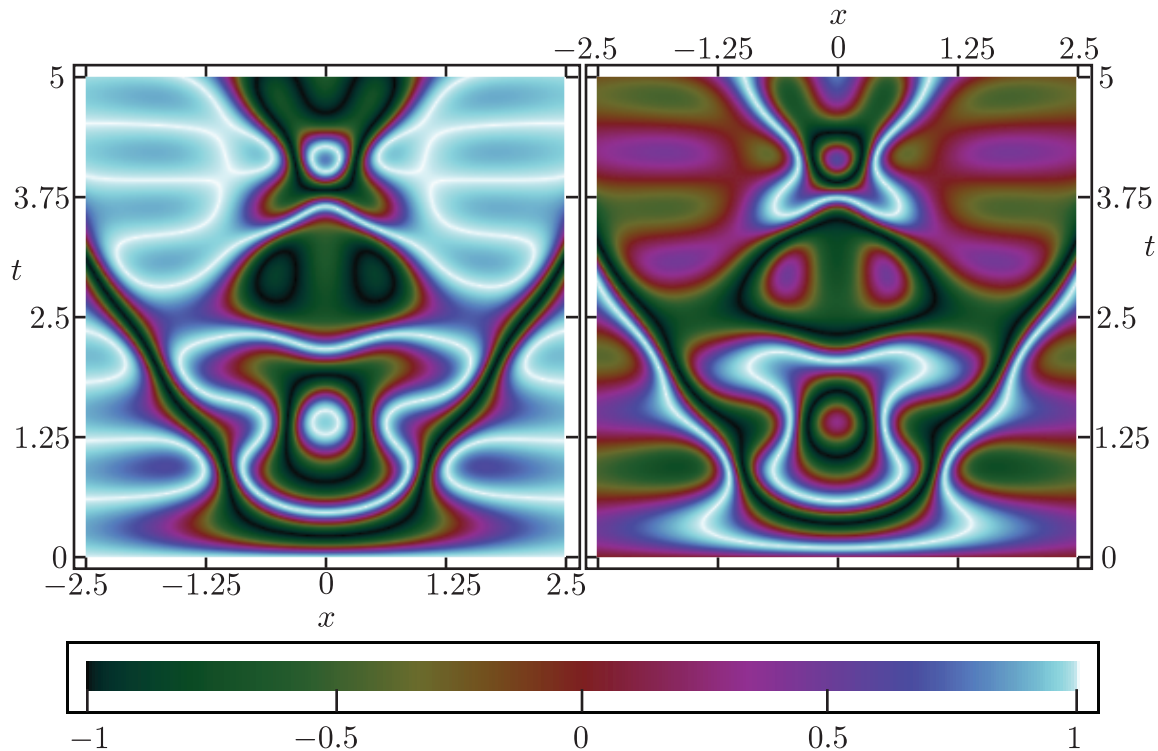


FIGURE 1.1. Density plots of  $\cos(u(x,t))$  (left) and  $\sin(u(x,t))$  (right) for the exact solution of the Cauchy problem (1.1) with  $\epsilon = 0.1875$  for  $F(x) \equiv 0$  and  $G(x) = -3 \operatorname{sech}(x)$ . For this data, separatrix crossings occur for  $x = \pm \operatorname{arcsech}(\frac{2}{3}) \approx \pm 0.962$ .

formula that is expressible in terms of modern special functions, specifically solutions of certain nonlinear ordinary differential equations of Painlevé type.

To formulate our results precisely requires that we first set up some background material; we hope that the reader will bear with us until §1.3 where the full details will be presented. In the meantime, we can describe the semiclassical asymptotics of  $u(x,t)$  in the region illustrated in Figure 1.5 by saying that this region contains a curvilinear network of isolated kink-type solutions of the sine-Gordon equation (with the approximation  $u(x,t) \approx \pi \pmod{2\pi}$  holding in between the kinks) lying along the graphs of certain rational solutions of inhomogeneous Painlevé-II equations. At spacetime points associated with poles and zeros of these solutions, the kinks collide in a grazing fashion as can be modeled by a more complicated exact solution of the sine-Gordon equation corresponding to a double soliton. The phenomenon of an asymptotically universal wave pattern being described by simple “solitonic” solutions located in space and time according to a rather more transcendental solution of a Painlevé-type equation was only recently observed for the first time in another problem by Bertola and Tovbis [5].

**1.1. Notation and terminology.** All power functions  $z^p$  will be assumed to be defined for nonintegral real  $p$  as the principal branch with branch cut  $z < 0$  and with  $-|p|\pi < \arg(z^p) < |p|\pi$ .

If  $G(\cdot)$  is a function of  $x$  for which the equation  $G(x) = -2$  has a unique solution for  $x > 0$ , then we denote this solution by  $x = x_{\text{crit}}$ . Let us use the term *criticality* to describe the point  $x = x_{\text{crit}}$  and  $t = 0$ , and *near criticality* to mean  $x \approx x_{\text{crit}}$  and  $t \approx 0$ .

For any rational function  $\mathcal{R} = \mathcal{R}(y)$ , let  $\mathcal{P}(\mathcal{R})$  denote the finite set of real poles of  $\mathcal{R}$  and let  $\mathcal{Z}(\mathcal{R})$  denote the finite set of real zeros of  $\mathcal{R}$ . Finally, for a finite set  $\mathcal{S} \subset \mathbb{R}$  and a real number  $y \in \mathbb{R}$  denote by

$$|y - \mathcal{S}| := \min_{y_0 \in \mathcal{S}} |y - y_0| \quad (1.3)$$

the distance between  $y$  and  $\mathcal{S}$ .

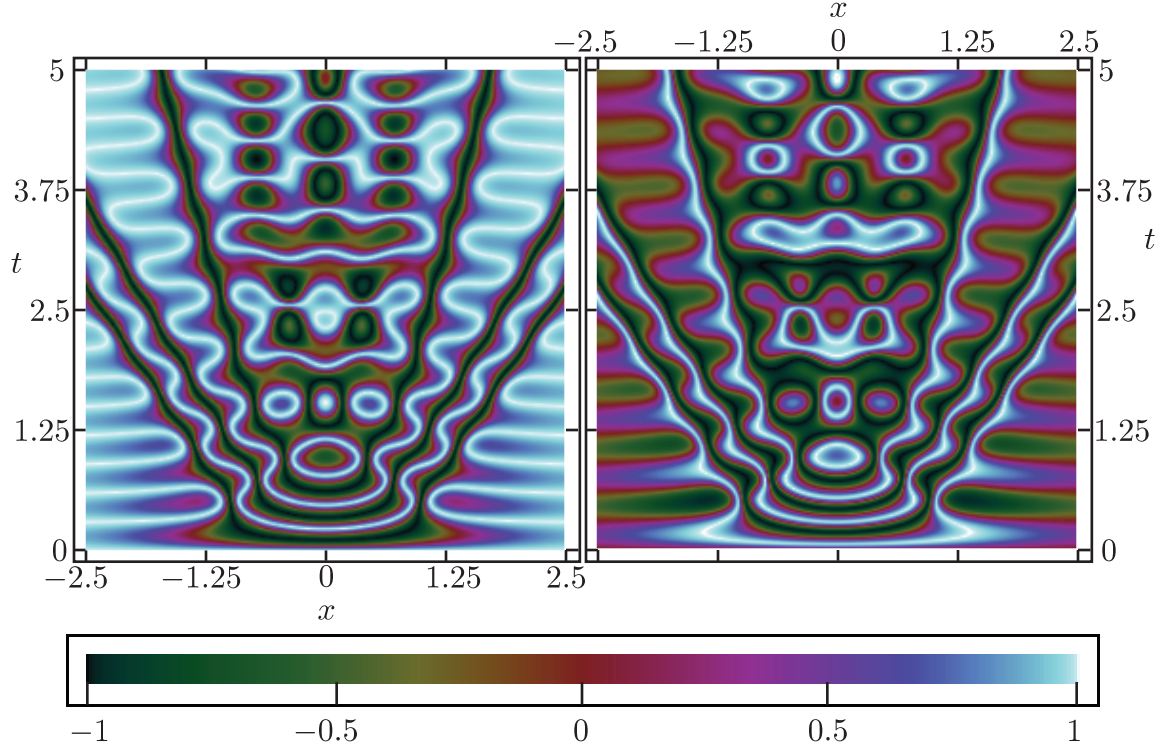


FIGURE 1.2. Same as Figure 1.1 but with  $\epsilon = 0.09375$ .

If  $\Sigma$  is an oriented contour in the complex plane and  $f$  is a function analytic in the complement of  $\Sigma$ , we will use subscripts  $f_+(\xi)$  and  $f_-(\xi)$  to refer to the boundary values taken by  $f(w)$  as  $w \rightarrow \xi \in \Sigma$  nontangentially from the left and right respectively.

We will make frequent use of the Pauli matrices:

$$\sigma_1 := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_2 := \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_3 := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \quad (1.4)$$

With the sole exception of these three, we write all matrices with boldface capital letters.

We use the Landau notation for most estimates, with “big-oh” written  $\mathcal{O}$  and “little-oh” written  $\mathfrak{o}$ . Also, if  $q_1, q_2, \dots, q_m$  are some quantities, we will use the shorthand notation  $\mathcal{O}(q_1, \dots, q_m)$  to represent a quantity that is bounded by a linear combination of  $|q_1|, \dots, |q_m|$ , that is,  $\mathcal{O}(q_1, \dots, q_m) = \mathcal{O}(q_1) + \dots + \mathcal{O}(q_m)$ .

We will be dealing with several matrix functions in which the matrix symbol carries subscripts and superscripts, so we will use a special notation for Taylor/Laurent expansion coefficients of such matrix functions: if  $\mathbf{M}_a^b(\tau)$  is such a matrix function and  $\tau_0$  is a point about which this matrix is to be expanded, we write the Taylor expansion in the form

$$\mathbf{M}_a^b(\tau) = \sum_{k=0}^{\infty} [\tau_0^k \mathbf{M}_a^b] (\tau - \tau_0)^k. \quad (1.5)$$

We also use this notation in the case that  $\tau_0 = \infty$  with obvious modifications.

**1.2. Assumptions and definition of fluxon condensates.** We study the Cauchy initial-value problem (1.1) under exactly the same assumptions used in our earlier work [7]:

**Assumption 1.1.** *The initial conditions for (1.1) satisfy  $F(x) \equiv 0$ .*

Assumption 1.1 asserts that the initial data is of pure-impulse type. This is important because it implies that the direct scattering problem for the solution of the Cauchy problem by the inverse-scattering transform

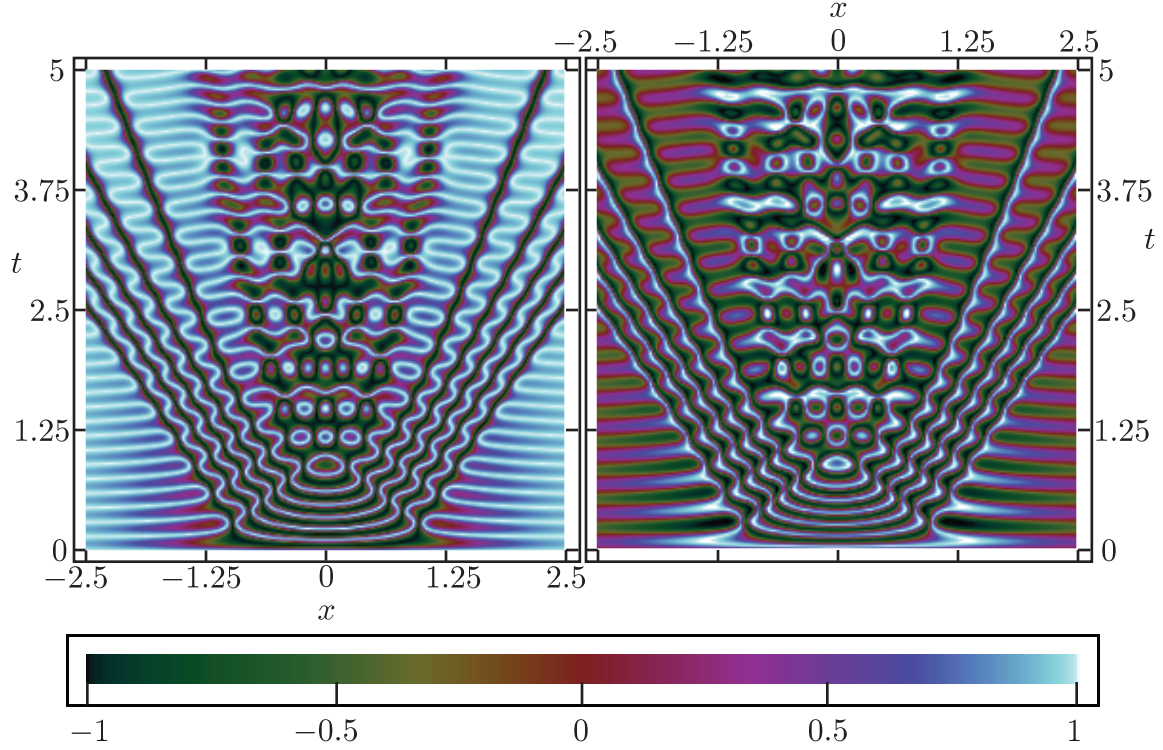


FIGURE 1.3. Same as Figure 1.1 but with  $\epsilon = 0.046875$ .

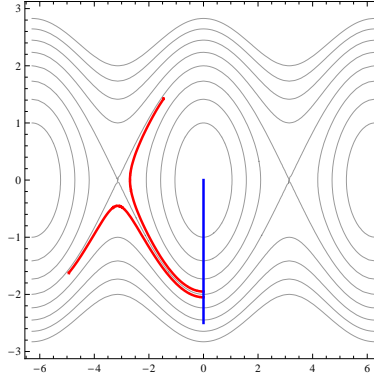


FIGURE 1.4. The initial data  $u(x, 0) = 0$  and  $u_T(x, 0) = G(x)$  with  $G(0) < -2$  plotted parametrically in the  $(u, u_T)$ -plane (blue). Orbits of the simple pendulum are shown with gray curves. Orbits corresponding to nearby values of  $x$  (red curves) can diverge after a descent to a neighborhood of a saddle point in finite  $T$ , if  $x \approx \pm x_{\text{crit}} = G^{-1}(-2)$ .

reduces from the Faddeev-Takhtajan eigenvalue problem to the better-understood nonselfadjoint Zakharov-Shabat eigenvalue problem.

**Assumption 1.2.** The function  $G$  is a nonpositive function of Klaus-Shaw type, that is,  $G \in L^1(\mathbb{R}) \cap C^1(\mathbb{R})$  and  $G$  has a unique local (and global) minimum.

As was shown by Klaus and Shaw in [21], Assumption 1.2 provides a useful and important confinement property of the discrete spectrum of the nonselfadjoint Zakharov-Shabat eigenvalue problem associated to



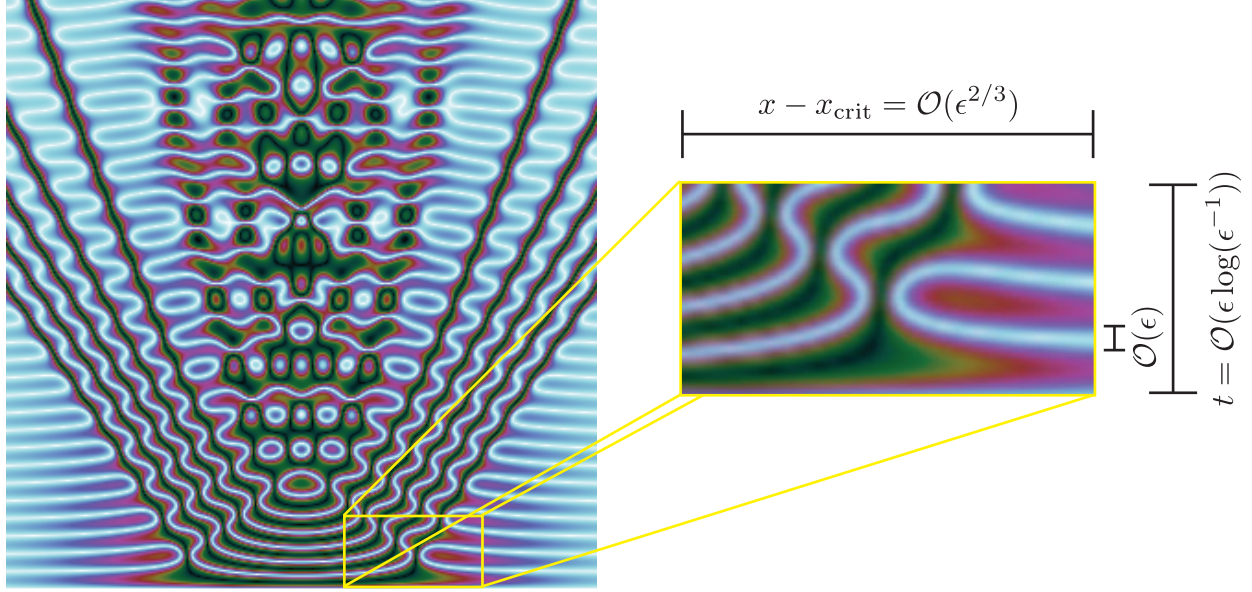


FIGURE 1.5. The region of the  $(x, t)$ -plane in which the fluxon condensate is analyzed.

the potential  $G$ . This allows us to formulate Riemann-Hilbert Problem 1.1 below on a system of contours very close to the unit circle and a transecting negative real interval, a set whose image under the function  $E(\cdot)$  defined in (1.10) below is a segment of the imaginary axis.

**Assumption 1.3.** The function  $G$  is even in  $x$ :  $G(-x) = G(x)$ , placing the unique minimum of  $G$  at  $x = 0$ .

Assumption 1.3 is admittedly less important (we believe that with appropriate modifications our results all hold true without it) but it allows for a substantial simplification of our analysis. We point out that under Assumptions 1.2 and 1.3 the function  $G$  restricted to  $\mathbb{R}_+$  has a well-defined inverse function  $G^{-1} : (G(0), 0) \rightarrow \mathbb{R}_+$ .

**Assumption 1.4.** The function  $G$  is strictly increasing and real-analytic at each  $x > 0$ , and the positive and real-analytic function

$$\mathcal{G}(m) := \frac{\sqrt{m}\sqrt{G(0)^2 - m}}{2G'(G^{-1}(-\sqrt{m}))}, \quad 0 < m < G(0)^2 \quad (1.6)$$

can be analytically continued to neighborhoods of  $m = 0$  and  $m = G(0)^2$  with  $\mathcal{G}(0) > 0$  and  $\mathcal{G}(G(0)^2) > 0$ .

The analyticity of  $G$  for  $x > 0$  and that of  $\mathcal{G}$  up to the endpoints of the interval  $[0, G(0)^2]$  as guaranteed by Assumption 1.4 are both absolutely crucial to our method of analysis. It is the analyticity of  $G$  that implies that of  $\Psi$  defined in (1.8) below, and hence of  $\theta_0$  defined by (1.10) and (1.14), and these are used to convert a “primordial” Riemann-Hilbert problem of inverse scattering that involves a large number (inversely proportional to  $\epsilon > 0$ ) of pole singularities into the simpler Riemann-Hilbert Problem 1.1 to be formulated below. The latter problem has no poles, but only jump discontinuities along contours, and hence is amenable to the Deift-Zhou steepest descent technique [16] of rigorous asymptotic analysis.

**Assumption 1.5.** The small number  $\epsilon$  lies in the infinite sequence

$$\epsilon = \epsilon_N := \frac{\|G\|_1}{4\pi N}, \quad N = 1, 2, 3, \dots \quad (1.7)$$

Assumption 1.5 is important because it minimizes the effect of *spectral singularities*, events occurring infinitely often as  $\epsilon \downarrow 0$  at which discrete eigenvalues are born from the continuous spectrum. When spectral

singularities occur, the reflection coefficient has poles in the continuous spectrum, and without Assumption 1.5 (or some suitable approximation thereof) the reflection coefficient cannot be neglected uniformly on the continuous spectrum. Assumption 1.5 is important because our approach is based on neglecting the reflection coefficient entirely. (It can be shown to be small in the semiclassical limit without Assumption 1.5 except near points where spectral singularities can occur.)

**Assumption 1.6.** *The function  $G(x)$  satisfies  $G(0) < -2$ .*

It is this last assumption that ensures that there exist exactly two points  $x = \pm x_{\text{crit}}$  at which the initial data curve  $(F, G) = (0, G(x))$  crosses the separatrix of the simple pendulum phase portrait. Therefore, Assumption 1.6 guarantees that the phenomenon we wish to study in this paper actually occurs for the initial data in question.

Given initial data for the Cauchy problem (1.1) satisfying these assumptions, we construct a sequence of exact solutions  $u = u_N(x, t)$  of the sine-Gordon equation  $\epsilon^2 u_{tt} - \epsilon^2 u_{xx} + \sin(u) = 0$  for  $\epsilon = \epsilon_N$ ,  $N = 1, 2, 3, \dots$  called the *fluxon condensate* associated with the given initial data. While  $u_N(x, t)$  is an exact solution of the partial differential equation for each  $N$ , it generally does not satisfy exactly the given initial conditions. However, it has been proved [7, Corollary 1.1] that  $u_N(x, 0) = \mathcal{O}(\epsilon_N)$  holds modulo  $4\pi$  and  $\epsilon_N u_{N,t}(x, 0) = G(x) + \mathcal{O}(\epsilon_N)$  where the error estimates are valid pointwise for  $x \neq 0$  and  $|x| \neq x_{\text{crit}}$ , and uniformly on compact subsets of the set of pointwise accuracy. We strengthen this convergence to include the points  $x$  near where  $|x| = x_{\text{crit}}$  in Theorem 1.1 below.

The fluxon condensate is constructed as follows. First one defines a function  $\Psi(\lambda)$  by setting

$$\Psi(\lambda) := \frac{1}{2} \int_0^{G^{-1}(-v)} \sqrt{G(s)^2 - v^2} ds, \quad \lambda = \frac{iv}{4}, \quad 0 < v < -G(0). \quad (1.8)$$

Note that  $\Psi$  is a strictly decreasing function of  $v$  where defined. This fact allows us to define a sequence of numbers  $\{\lambda_{N,k}^0\}_{k=0}^{N-1}$  by solving the equation

$$\Psi(\lambda_{N,k}^0) = \pi \epsilon_N \left( k + \frac{1}{2} \right), \quad k = 0, 1, 2, \dots, N-1. \quad (1.9)$$

(These positive imaginary numbers are approximate eigenvalues for the Zakharov-Shabat eigenvalue problem associated with the Klaus-Shaw potential  $G$ , and (1.9) is a kind of *Bohr-Sommerfeld quantization rule* for that nonselfadjoint problem.) Setting

$$E(w) := \frac{i}{4} \left[ (-w)^{1/2} + (-w)^{-1/2} \right] \quad \text{and} \quad D(w) := \frac{i}{4} \left[ (-w)^{1/2} - (-w)^{-1/2} \right], \quad |\arg(-w)| < \pi, \quad (1.10)$$

we define

$$Q(w; x, t) := E(w)x + D(w)t, \quad |\arg(-w)| < \pi, \quad (1.11)$$

and

$$\Pi_N(w) := \prod_{k=0}^{N-1} \frac{E(w) + \lambda_{N,k}^0}{E(w) - \lambda_{N,k}^0}, \quad |\arg(-w)| < \pi. \quad (1.12)$$

This function is meromorphic and has no zeros where defined. It has  $2N$  poles (counted with multiplicity), the set of which we denote as  $P_N$ . Generically (with respect to deformations of  $G$ ) the poles are all simple, and there are  $2N_{\text{B}}$  of them in complex-conjugate pairs on the unit circle in the  $w$ -plane, and  $2N_{\text{K}}$  of them in pairs on the negative real axis in involution with respect to the map  $w \mapsto 1/w$ . As  $N \uparrow \infty$  (or  $\epsilon_N \downarrow 0$ ) the poles accumulate on the whole unit circle and the interval  $[\mathfrak{a}, \mathfrak{b}]$ , where

$$\mathfrak{a} := -\frac{1}{4} \left( \sqrt{G(0)^2 - 4} - G(0) \right)^2, \quad \mathfrak{b} := -\frac{1}{4} \left( \sqrt{G(0)^2 - 4} + G(0) \right)^2 = \frac{1}{\mathfrak{a}}. \quad (1.13)$$

Clearly, both  $\mathfrak{a}$  and  $\mathfrak{b}$  are independent of  $N$ , and  $\mathfrak{a} < -1 < \mathfrak{b} < 0$ . Were Assumption 1.6 not satisfied, we would have  $N_{\text{K}} = 0$ , and the poles of  $\Pi_N$  would accumulate only in a complex-conjugate symmetric incomplete arc of the unit circle containing the point  $w = 1$ . In nongeneric cases it can happen that there is a double pole at  $w = -1$ . We consider the accumulation locus of poles,  $P_\infty := [\mathfrak{a}, \mathfrak{b}] \cup S^1$  to be an oriented contour, with orientation of the two intervals  $(\mathfrak{a}, -1)$  and  $(-1, \mathfrak{b})$  both toward  $w = -1$ , and with orientation of the upper and lower semicircles from  $w = -1$  toward  $w = 1$ . See Figure 1.6. We also use the notation

$$\theta_0(w) := \Psi(E(w)). \quad (1.14)$$

Then, define a function  $L(w)$  by a Cauchy integral:

$$L(w) := \frac{(-w)^{1/2}}{\pi} \int_{P_\infty} \frac{\theta_0(y)}{(-y)^{1/2} y - w} dy, \quad w \in \mathbb{C} \setminus (P_\infty \cup \mathbb{R}_+), \quad (1.15)$$

and then set

$$Y_N(w) := \Pi_N(w) e^{-L(w)/\epsilon_N}, \quad w \in \mathbb{C} \setminus (P_\infty \cup \mathbb{R}_+). \quad (1.16)$$

Finally, writing

$$\bar{L}(\xi) := \frac{1}{2}(L_+(\xi) + L_-(\xi)), \quad \xi \in P_\infty, \quad (1.17)$$

a function that turns out to have a well-defined analytic continuation to a full neighborhood of the self-intersection point of  $w = -1$  of the contour  $P_\infty$  (upon continuation from any of the four intersecting arcs), we set

$$T_N(\xi) := 2\Pi_N(\xi) \cos(\epsilon_N^{-1} \theta_0(\xi)) e^{-\bar{L}(\xi)/\epsilon_N}, \quad \xi \in P_\infty. \quad (1.18)$$

It can be shown [7, Proposition 3.1] that  $Y_N(w) = 1 + \mathcal{O}(\epsilon_N)$  holds uniformly on compact subsets of the open domain of definition, and that  $Y_N(w) = \mathcal{O}(1)$  holds if  $w \rightarrow \mathfrak{a}$  or  $w \rightarrow \mathfrak{b}$  nontangentially to the real axis. Similarly,  $T_N(\xi)$  extends from  $P_\infty$  to an analytic function  $T_N(w)$  on a neighborhood of  $w = -1$ , and  $T_N(w) = 1 + \mathcal{O}(\epsilon_N)$  holds uniformly on this neighborhood as well as on  $P_\infty$ , as long as  $w$  is bounded away from  $w = \mathfrak{a}$  and  $w = \mathfrak{b}$ ; however the estimate  $T_N(\xi) = \mathcal{O}(1)$  holds uniformly for  $\xi \in P_\infty$ .

Consider the contour  $\Sigma_N$  illustrated in Figure 1.6. Let  $g(w)$  denote a function analytic in the domain

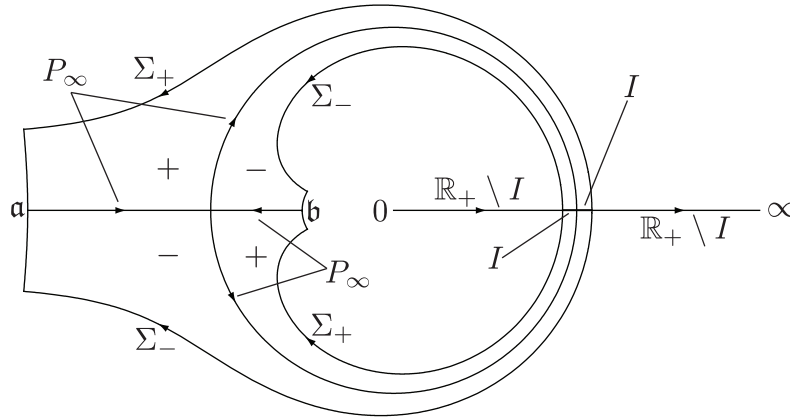


FIGURE 1.6. The contour  $\Sigma_N$  of discontinuity for the sectionally analytic function  $N(w)$ . With the exception of the two components of  $\mathbb{R}_+ \setminus I$  which are oriented left-to-right, the contour arcs are oriented with the regions labeled “+” on the left, and with the regions labeled “-” on the right. The image of  $\Sigma_N \setminus \mathbb{R}_+$  under  $w \mapsto E(w)$  with  $E$  defined by (1.10) consists of the four straight line segments: (i)  $\Re\{E\} = -\delta$ ,  $0 < \Im\{E\} \leq -G(0)/4$ , (ii)  $\Re\{E\} = 0$ ,  $0 < \Im\{E\} \leq -G(0)/4$ , (iii)  $\Re\{E\} = \delta$ ,  $0 < \Im\{E\} \leq -G(0)/4$ , and (iv)  $|\Re\{E\}| \leq \delta$ ,  $\Im\{E\} = -G(0)/4$ . Here  $\delta > 0$  is a sufficiently small number. The self-intersection point  $w = -1$  is the critical point of  $E$ .

$w \in \mathbb{C} \setminus (P_\infty \cup \mathbb{R}_+)$ , satisfying the symmetry  $g(w^*) = g(w)^*$  and the conditions that  $g(0) = 0$  and  $g(w) \rightarrow 0$  as  $w \rightarrow \infty$ . We assume also that  $g$  takes well-defined boundary values on  $P_\infty \cup \mathbb{R}_+$  in the classical sense (Hölder continuity up to the boundary), and that the boundary values satisfy  $g_+(\xi) + g_-(\xi) = 0$  for  $\xi \in \mathbb{R}_+$ . Set

$$\theta(\xi) := -i(g_+(\xi) - g_-(\xi)) \quad \text{and} \quad \phi(\xi; x, t) := 2iQ(\xi; x, t) + \bar{L}(\xi) - g_+(\xi) - g_-(\xi), \quad \xi \in P_\infty. \quad (1.19)$$

Then, to determine the fluxon condensate, one solves the following Riemann-Hilbert problem.

**Riemann-Hilbert Problem 1.1** (Basic Problem of Inverse Scattering). *Given  $g$  as above, find a  $2 \times 2$  matrix function  $N(w) = N_N^g(w; x, t)$  of the complex variable  $w$  with the following properties:*



**Analyticity:**  $\mathbf{N}(w)$  is analytic for  $w \in \mathbb{C} \setminus \Sigma_{\mathbf{N}}$ .

**Jump condition:**  $\mathbf{N}(w)$  is Hölder continuous up to  $\Sigma_{\mathbf{N}}$ . The boundary values  $\mathbf{N}_{\pm}(\xi)$  taken by  $\mathbf{N}(w)$  on the various arcs of  $\Sigma_{\mathbf{N}}$  are required to satisfy the following jump conditions

$$\mathbf{N}_+(\xi) = \sigma_2 \mathbf{N}_-(\xi) \sigma_2, \quad \xi \in \mathbb{R}_+ \setminus I, \quad (1.20)$$

$$\mathbf{N}_+(\xi) = \sigma_2 \mathbf{N}_-(\xi) \sigma_2 \begin{bmatrix} 1 + e^{i(\theta_{0+}(\xi) - \theta_{0-}(\xi))/\epsilon_N} & iY_{N-}(\xi) e^{-i\theta_{0-}(\xi)/\epsilon_N} e^{-F(\xi; x, t)/\epsilon_N} \\ -Y_{N+}(\xi) e^{i\theta_{0+}(\xi)/\epsilon_N} e^{F(\xi; x, t)/\epsilon_N} & 1 \end{bmatrix}, \quad \xi \in I, \quad (1.21)$$

where  $F(\xi; x, t) := 2iQ_+(\xi; x, t) + L_+(\xi) - 2g_+(\xi)$ ,

$$\mathbf{N}_+(\xi) = \mathbf{N}_-(\xi) \begin{bmatrix} e^{-i\theta(\xi)/\epsilon_N} & 0 \\ -iT_N(\xi) e^{\phi(\xi; x, t)/\epsilon_N} & e^{i\theta(\xi)/\epsilon_N} \end{bmatrix}, \quad \xi \in P_{\infty}, \quad (1.22)$$

and,

$$\mathbf{N}_+(\xi) = \mathbf{N}_-(\xi) \begin{bmatrix} 1 & 0 \\ -iY_N(\xi) e^{(2iQ(\xi; x, t) + L(\xi) \pm i\theta_0(\xi) - 2g(\xi))/\epsilon_N} & 1 \end{bmatrix}, \quad \xi \in \Sigma_{\pm}. \quad (1.23)$$

**Normalization:** The following normalization condition holds:

$$\lim_{w \rightarrow \infty} \mathbf{N}(w) = \mathbb{I}, \quad (1.24)$$

where the limit is uniform with respect to angle for  $|\arg(-w)| < \pi$ .

It turns out that for each choice of the function  $g$ , each  $N = 1, 2, 3, \dots$ , and each  $(x, t) \in \mathbb{R}^2$  there exists a unique solution of this Riemann-Hilbert problem. Moreover, the product  $\mathbf{M}(w) = \mathbf{M}_N(w; x, t) := \mathbf{N}_N^g(w; x, t) e^{g(w)\sigma_3/\epsilon_N}$  does not depend on the choice of the function  $g$ . The solution  $\mathbf{N}_N^g(w; x, t)$  has convergent expansions near  $w = 0$  and  $w = \infty$  of the form

$$\mathbf{N}_N^g(w; x, t) = \mathbb{I} + \sum_{k=1}^{\infty} [\mathbf{N}_N^g]_k(x, t) (-w)^{-k/2}, \quad \text{for } |w| \text{ sufficiently large}, \quad (1.25)$$

and

$$\mathbf{N}_N^g(w; x, t) = \sum_{k=0}^{\infty} [\mathbf{N}_N^g]_k(x, t) (-w)^{k/2}, \quad \text{for } |w| \text{ sufficiently small}. \quad (1.26)$$

**Definition 1.1** (Fluxon condensates). *Given a function  $G$  satisfying Assumptions 1.1 through 1.6, the fluxon condensate associated with  $G$  is the family of functions  $\{u_N(x, t)\}_{N=1}^{\infty}$  defined modulo  $4\pi$  by the equations*

$$\cos(\tfrac{1}{2}u_N(x, t)) := [\mathbf{N}_N^g]_0(x, t)_{11} \quad \text{and} \quad \sin(\tfrac{1}{2}u_N(x, t)) := [\mathbf{N}_N^g]_0(x, t)_{21}. \quad (1.27)$$

Note that these are independent of  $g$  because  $g(0) = 0$ .

While it is possible to extract formulae for derivatives of  $u_N(x, t)$  with respect to  $x$  and  $t$  by the chain rule, it is preferable to have formulae that do not require differentiation of  $\mathbf{N}_N^g(w; x, t)$ , as this will require control of derivatives of error terms. However, it can be shown also that the following formula holds:

$$\epsilon_N \frac{\partial u_N}{\partial t}(x, t) = [\mathbf{N}_N^g]_1(x, t)_{12} + [\mathbf{N}_N^g]_0(x, t)^{-1} [\mathbf{N}_N^g]_1(x, t)_{12}, \quad (1.28)$$

and this does not require differentiation of  $\mathbf{N}_N^g(w; x, t)$  with respect to  $t$  (it also turns out to be independent of choice of  $g$ ). Each function  $u_N(x, t)$  of the fluxon condensate is an exact solution of the sine-Gordon equation with  $\epsilon = \epsilon_N$ , and  $u_N(-x, t) = u_N(x, t)$ .

While Riemann-Hilbert Problem 1.1 is the most convenient starting point for our local analysis near criticality, it is not the most fundamental representation of  $u_N(x, t)$ . Indeed,  $u_N(x, t)$  is a *reflectionless potential* for the direct scattering problem of the Lax pair for the sine-Gordon equation, and this means that it can be obtained from a purely “discrete” Riemann-Hilbert problem whose solution is a rational function on the Riemann surface of  $y^2 = -w$ , with poles on both sheets over the points of  $P_N$ . Since the number of poles is increasing with  $N$ , some preparations are required to recast the problem in a suitable form for addressing the limit  $N \rightarrow \infty$ . These preparations are detailed in [7], and they take the form of a sequence of explicit transformations resulting in the equivalent Riemann-Hilbert problem 1.1. In general, a number of choices are made along the way because the transformations that are required to enable the subsequent

asymptotic analysis depend on  $(x, t)$ ; however for  $x \approx x_{\text{crit}} > 0$  and  $t$  small only the simplest of the choices detailed in [7] are required. For readers familiar with the terminology of that paper, we are assuming that  $\Delta = \emptyset$ , and correspondingly,  $Y_N(w)$  is the function called  $Y^\nabla(w)$  in [7] while  $T_N(w)$  is the function called  $T^\nabla(w)$  in [7]. Also, for our purposes we need make no distinction between the contour called  $\Sigma^\nabla \cup \Sigma^\Delta$  (which would just be  $\Sigma^\nabla$  in the case that  $\Delta = \emptyset$ ) and the contour  $P_\infty$ , and this implies that the set  $Z \subset \mathbb{C}$  that lies between these two contours is empty. Finally, to derive Riemann-Hilbert Problem 1.1 from the results of [7] we used the facts that for  $\xi \in I$ ,  $Q_+(\xi; x, t) + Q_-(\xi; x, t) = L_+(\xi) + L_-(\xi) = 0$ .

The solution of the sine-Gordon equation by inverse-scattering methods is a subject with a long history. When the sine-Gordon equation is written in terms of characteristic coordinates  $x \pm t$  it fits naturally into the hierarchy of the Ablowitz-Kaup-Newell-Segur or Zakharov-Shabat scattering problem. The *characteristic* Cauchy problem for the sine-Gordon equation was integrated in this framework by Ablowitz, Kaup, Newell, and Segur [1] and by Zakharov, Takhtajan, and Faddeev [25]. The more physically-relevant problem of the Cauchy problem in laboratory coordinates as posed in (1.1) required new methodology, and the solution of this Cauchy problem by the inverse scattering method was first outlined by Kaup [20]. An account of the solution of the Cauchy problem in laboratory coordinates is given in the textbook of Faddeev and Takhtajan [17]. Some further analytical details needed to make the theory completely rigorous were supplied by Zhou [26] and later by Cheng, Venakides, and Zhou [8]. A self-contained account of the Riemann-Hilbert formulation of the inverse-scattering solution of the Cauchy problem (1.1) assuming only that at each instant of time the solution has  $L^1$ -Sobolev regularity can be found in our paper [6, Appendix A], and a direct proof that the sine-Gordon equation preserves this degree of regularity if it is initially present is given in [6, Appendix B]. In our recent paper [7] we found that to describe the semiclassical limit for the Cauchy problem (1.1) it is useful to reformulate the Riemann-Hilbert problem in the complex plane of a square root of the spectral variable used in [6, Appendix A]; this makes it easier to express the asymptotic solutions in terms of Riemann theta functions of the lowest possible genus. As we view the current paper as a continuation of our work in [7] we use the same formulation here.

**1.3. Statement of results.** Let  $x_{\text{crit}}$  be defined by

$$x_{\text{crit}} := G^{-1}(-2) > 0, \quad (1.29)$$

and define a positive constant  $\nu$  by

$$\nu := \frac{1}{12G'(x_{\text{crit}})} > 0. \quad (1.30)$$

Also, set

$$\Delta x := x - x_{\text{crit}}. \quad (1.31)$$

All of our results concern the asymptotic behavior of the fluxon condensate  $u_N(x, t)$  in the small region near criticality where  $t = \mathcal{O}(\epsilon_N \log(\epsilon_N^{-1}))$  and  $\Delta x = \mathcal{O}(\epsilon_N^{2/3})$  as shown in Figure 1.5.

Our first result is concerned with the relevance of the fluxon condensate associated with  $G(x)$  to the Cauchy initial-value problem.

**Theorem 1.1** (Initial accuracy of fluxon condensates). *Suppose that  $x \pm x_{\text{crit}} = \mathcal{O}(\epsilon_N^{2/3})$ . Then uniformly for such  $x$ ,*

$$u_N(x, 0) = \mathcal{O}(\epsilon_N^{1/3}) \pmod{4\pi} \quad \text{and} \quad \epsilon_N \frac{\partial u_N}{\partial t}(x, 0) = G(x) + \mathcal{O}(\epsilon_N^{1/3}). \quad (1.32)$$

This result extends that of [7, Corollary 1.1] to suitable neighborhoods of the points  $x = \pm x_{\text{crit}}$ .

**Theorem 1.2** (Main approximation theorem). *There exist multiscale approximating functions  $\dot{C}(x, t; \epsilon_N)$  and  $\dot{S}(x, t; \epsilon_N)$  (defined in detail in §7.3, and depending on initial data only through the constant  $\nu > 0$ ) such that*

$$\begin{aligned} \cos(\tfrac{1}{2}u_N(x, t)) &= \dot{C}(x, t; \epsilon_N) + \mathcal{O}(\epsilon_N^{1/6}) \\ \sin(\tfrac{1}{2}u_N(x, t)) &= \dot{S}(x, t; \epsilon_N) + \mathcal{O}(\epsilon_N^{1/6}) \end{aligned} \quad (1.33)$$

*with the error terms being uniform for  $\Delta x = \mathcal{O}(\epsilon_N^{2/3})$  and  $t = \mathcal{O}(\epsilon_N \log(\epsilon_N^{-1}))$ .*

In fact, we will really show that the error term is significantly smaller, namely  $\mathcal{O}(\epsilon_N^{1/3})$ , over most of the small region of the  $(x, t)$ -plane where the above Theorem provides an asymptotic description of the dispersive breakup of the pendulum separatrix. Now, the multiscale model provided by Theorem 1.2 serves to establish universality of the behavior near the critical point, but the formulae for  $\dot{C}(x, t; \epsilon_N)$  and  $\dot{S}(x, t; \epsilon_N)$  are rather complicated, so it is useful to give some more detailed information by focusing on smaller parts of the  $(x, t)$ -plane near criticality.

To render our results in a more elementary fashion, we need to first define a certain hierarchy of rational functions. First define

$$\mathcal{U}_0(z) := 1 \quad \text{and} \quad \mathcal{V}_0(z) := -\frac{1}{6}z. \quad (1.34)$$

Then define  $\{(\mathcal{U}_m, \mathcal{V}_m)\}_{m \in \mathbb{Z}}$  by the forward recursion

$$\mathcal{U}_{m+1}(z) := -\frac{1}{6}z\mathcal{U}_m(z) - \frac{\mathcal{U}_m'(z)^2}{\mathcal{U}_m(z)} + \frac{1}{2}\mathcal{U}_m''(z) \quad \text{and} \quad \mathcal{V}_{m+1}(z) := \frac{1}{\mathcal{U}_m(z)} \quad (1.35)$$

and the backward recursion

$$\mathcal{U}_{m-1}(z) := \frac{1}{\mathcal{V}_m(z)} \quad \text{and} \quad \mathcal{V}_{m-1}(z) := \frac{1}{2}\mathcal{V}_m''(z) - \frac{\mathcal{V}_m'(z)^2}{\mathcal{V}_m(z)} - \frac{1}{6}z\mathcal{V}_m(z). \quad (1.36)$$

Up to constant factors,  $\mathcal{U}_m(z)$  and  $\mathcal{V}_m(z)$  are ratios of consecutive *Yablonskii-Vorob'ev* polynomials (see [13, 14]) and logarithmic derivatives of  $\mathcal{U}_m(z)$  and  $\mathcal{V}_m(z)$  are the unique [22] rational solutions of the inhomogeneous Painlevé-II equations. These observations are not necessary for us to state our results, and we will make further comments later at an appropriate point in the paper. The following proposition characterizes the behavior of these rational functions for large  $z$ .

**Proposition 1.1.** *For each  $m \in \mathbb{Z}$ ,*

$$\mathcal{U}_m(z) = \left(-\frac{z}{6}\right)^m (1 + \mathcal{O}(z^{-1})) \quad \text{and} \quad \mathcal{V}_m(z) = \left(-\frac{z}{6}\right)^{1-m} (1 + \mathcal{O}(z^{-1})) \quad (1.37)$$

as  $z \rightarrow \infty$ . In particular,  $\text{sgn}(\mathcal{U}_m(z)) = \text{sgn}(\mathcal{V}_m(z)) = 1$  for sufficiently large negative  $z$  while  $\text{sgn}(\mathcal{U}_m(z)) = -\text{sgn}(\mathcal{V}_m(z)) = (-1)^m$  for sufficiently large positive  $z$ .

*Proof.* It is obvious that the recursions (1.35) and (1.36) preserve rationality of the input functions  $(\mathcal{U}_m, \mathcal{V}_m)$ , so for each  $m \in \mathbb{Z}$ ,  $\mathcal{U}_m$  is a rational function of  $z$ . Therefore,  $\mathcal{U}_m$  has a Laurent expansion for large  $z \in \mathbb{C}$  of the form

$$\mathcal{U}_m(z) = c_m z^{k_m} (1 + \mathcal{O}(z^{-1})) \quad (1.38)$$

where  $c_m$  is a complex constant and  $k_m \in \mathbb{Z}$  is an exponent to be determined, and moreover this expansion is differentiable any number of times with respect to  $z$ . In particular, it follows that

$$\frac{\mathcal{U}_m'(z)^2}{\mathcal{U}_m(z)} = k_m^2 c_m z^{k_m-2} (1 + \mathcal{O}(z^{-1})) \quad \text{and} \quad \mathcal{U}_m''(z) = k_m(k_m-1)c_m z^{k_m-2} (1 + \mathcal{O}(z^{-1})) \quad (1.39)$$

as  $z \rightarrow \infty$ . This shows that the final two terms on the right-hand side of the formula for  $\mathcal{U}_{m+1}(z)$  given in (1.35) are subdominant compared to the term  $-z\mathcal{U}_m(z)/6$ , and therefore,

$$\mathcal{U}_{m+1}(z) = -\frac{z}{6}c_m z^{k_m} (1 + \mathcal{O}(z^{-1})) = -\frac{z}{6}\mathcal{U}_m(z) (1 + \mathcal{O}(z^{-1})), \quad z \rightarrow \infty. \quad (1.40)$$

This formula gives the recurrence relation for the large- $z$  asymptotics of  $\mathcal{U}_m$ . The asymptotic formula for  $\mathcal{U}_m(z)$  given in (1.37) follows by solving the recurrence with the base case  $\mathcal{U}_0(z) = 1$ , and that for  $\mathcal{V}_m(z)$  follows from the identity  $\mathcal{V}_m(z) = 1/\mathcal{U}_{m-1}(z)$ .  $\square$

**Theorem 1.3** (Superluminal kink asymptotics). *Fix an integer  $m$ , and suppose that  $\Delta x = \mathcal{O}(\epsilon_N^{2/3})$  while  $|t - \frac{2}{3}m\epsilon_N \log(\epsilon_N^{-1})| \leq \frac{1}{3}\epsilon_N \log(\epsilon_N^{-1})$ . Then*

$$\begin{aligned} \cos(\tfrac{1}{2}u_N(x, t)) &= (-1)^m \text{sgn}(\mathcal{U}_m(z)) \text{sech}(T_K) + E_{\cos}(x, t; \epsilon_N) \\ \sin(\tfrac{1}{2}u_N(x, t)) &= (-1)^{m+1} \tanh(T_K) + E_{\sin}(x, t; \epsilon_N), \end{aligned} \quad (1.41)$$

where the phase is

$$T_K := \frac{t}{\epsilon_N} - 2m \log \left( \frac{4\nu^{1/3}}{\epsilon_N^{1/3}} \right) + \log |\mathcal{U}_m(z)|, \quad (1.42)$$

and a rescaled spatial coordinate is given by

$$z := \frac{\Delta x}{2\nu^{1/3}\epsilon_N^{2/3}}. \quad (1.43)$$

The error terms satisfy

$$\lim_{\epsilon_N \downarrow 0} E_{\cos}(x, t; \epsilon_N) = \lim_{\epsilon_N \downarrow 0} E_{\sin}(x, t; \epsilon_N) = 0 \quad (1.44)$$

unless (i) both  $\epsilon_N^{-2m/3}e^{-t/\epsilon_N} = \mathcal{O}(\epsilon_N^{1/3})$  and  $|z - \mathcal{Z}(\mathcal{U}_m)| = \mathcal{O}(\epsilon_N^{1/3})$  or (ii) both  $\epsilon_N^{2m/3}e^{t/\epsilon_N} = \mathcal{O}(\epsilon_N^{1/3})$  and  $|z - \mathcal{P}(\mathcal{U}_m)| = \mathcal{O}(\epsilon_N^{1/3})$ . Moreover, given any interval  $[z_-, z_+]$  on which  $\log |\mathcal{U}_m(z)|$  is a bounded function, we have

$$E_{\cos}(x, t; \epsilon_N) = \mathcal{O}(\epsilon_N^{1/3}) \quad \text{and} \quad E_{\sin}(x, t; \epsilon_N) = \mathcal{O}(\epsilon_N^{1/3}) \quad (1.45)$$

holding uniformly for  $z \in [z_-, z_+]$  and  $|t - \frac{2}{3}m\epsilon_N \log(\epsilon_N^{-1})| \leq \frac{1}{3}\epsilon_N \log(\epsilon_N^{-1})$ .

Wherever the error terms vanish in the limit, it follows that  $\cos(u_N(x, t)) \rightarrow 2 \operatorname{sech}^2(T_K) - 1$  and that  $\sin(u_N(x, t)) \rightarrow -2\sigma \operatorname{sech}(T_K) \tanh(T_K)$ , where  $\sigma := \operatorname{sgn}(\mathcal{U}_m(z))$ . If we define a function  $u(T)$  modulo  $2\pi$  by

$$\begin{aligned} \cos(u(T)) &:= 2 \operatorname{sech}^2(T) - 1 \\ \sin(u(T)) &:= -2\sigma \operatorname{sech}(T) \tanh(T), \end{aligned} \quad (1.46)$$

then it is easy to check that  $u(T)$  is an  $X$ -independent solution of the unscaled sine-Gordon equation

$$\frac{\partial^2 u}{\partial T^2} - \frac{\partial^2 u}{\partial X^2} + \sin(u) = 0. \quad (1.47)$$

This exact solution represents a superluminal (infinite velocity) kink with unit magnitude *topological charge*

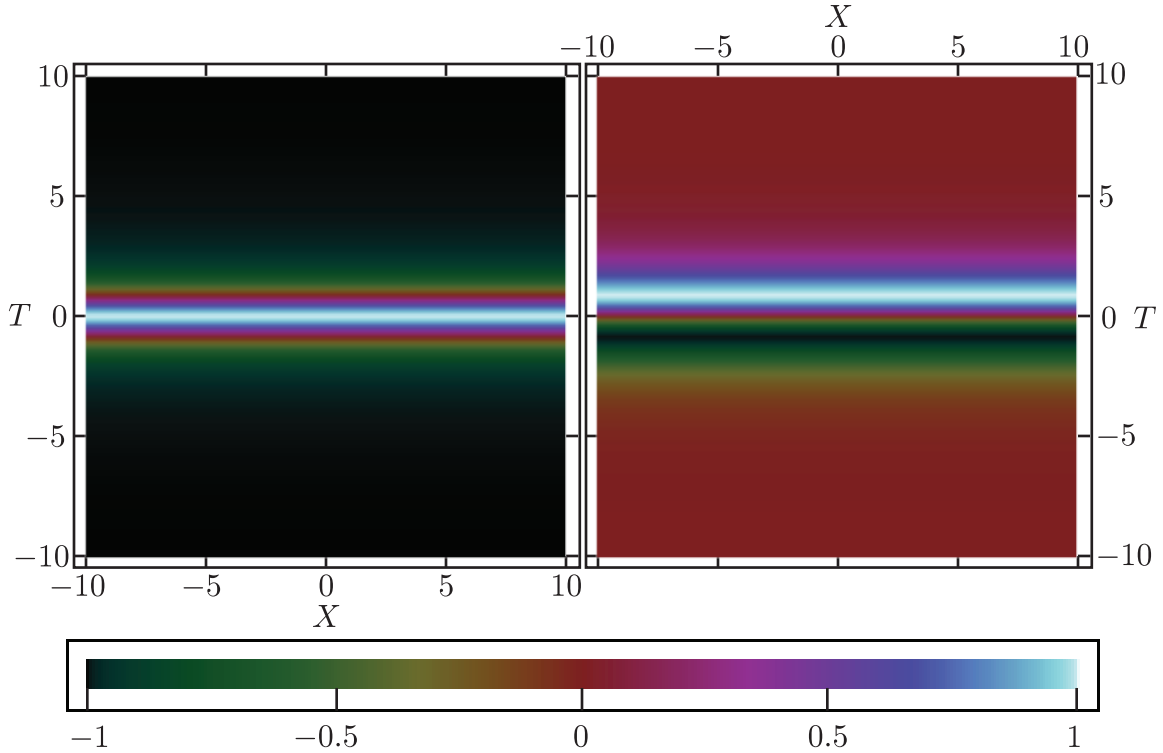


FIGURE 1.7. (Superluminal antikink). The cosine (left) and sine (right) of  $u(X, T) = u(T)$  defined by (1.46) with  $\sigma = 1$ . For a superluminal kink ( $\sigma = -1$ ), the plots are simply reflected through the horizontal  $X$ -axis.

$\sigma$ ; if  $\sigma = 1$  (respectively  $\sigma = -1$ ) then  $u$  decreases (respectively increases) by  $2\pi$  as  $T$  varies from  $T = -\infty$

to  $T = +\infty$ . Sometimes one refers to  $u(T)$  as a kink in the case  $\sigma = -1$  and as an *antikink* in the case  $\sigma = 1$ . Another way to describe  $u(T)$  is to say that it is a solution of the simple pendulum equation  $u'' + \sin(u) = 0$  that is homoclinic to the unstable equilibrium of a stationary inverted pendulum. See Figure 1.7.

The dominant part of the phase variable  $T_K$  defined by (1.42) is indeed a recentering and rescaling by  $\epsilon_N$  of  $t$ ; however  $T_K$  also contains weak  $x$ -dependence through the function  $\log |\mathcal{U}_m(z)|$  where  $z$  is proportional to  $x - x_{\text{crit}}$ . Therefore, the approximation of  $u_N(x, t)$  described in Theorem 1.3 is not an exact kink, but rather is one that is slowly modulated in the direction parallel to the wavefront. Indeed, the center of the approximating kink (where the pendulum angle is zero) corresponds to  $T_K = 0$ , which is a curve in the  $(x, t)$ -plane that is a scaled and translated version of the graph of the function  $-\log |\mathcal{U}_m(z)|$ . See Figure 1.8. An additional important observation is that the period of each approximating kink is proportional to  $\epsilon_N$ , but the strip in which it lives (indexed by the integer  $m$ ) corresponds to a time interval of length proportional to  $\epsilon_N \log(\epsilon_N^{-1})$ ; hence the kinks are in reality widely separated from each other, and therefore in “most” of the domain  $\Delta x = \mathcal{O}(\epsilon_N^{2/3})$  and  $t = \mathcal{O}(\epsilon_N \log(\epsilon_N^{-1}))$  covered by Theorem 1.3 it is fair to say that the approximate formula  $u_N(x, t) \approx \pi \pmod{2\pi}$  holds, which in the context of the coupled pendulum interpretation of sine-Gordon means that “most” of the pendula are approximately in the unstable inverted configuration.

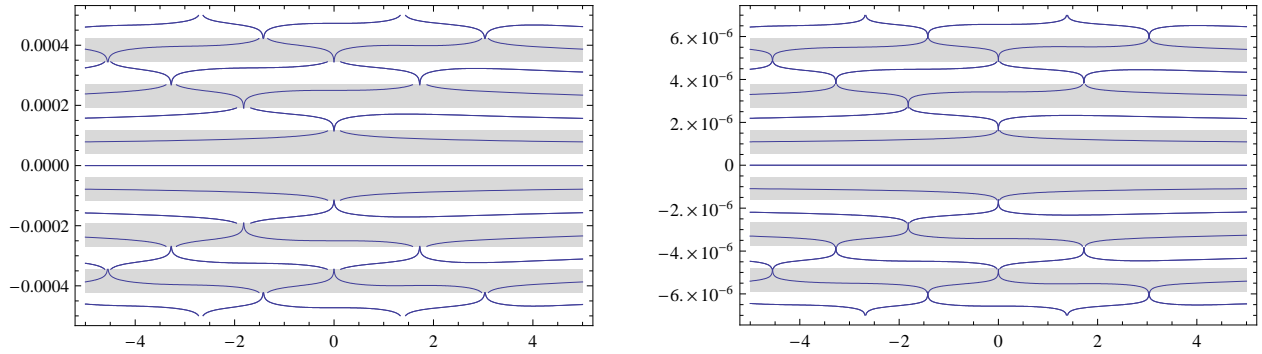


FIGURE 1.8. The horizontal strips in the  $(z, t)$ -plane in each of which Theorem 1.3 provides asymptotics for  $u_N(x, t)$ , shown with alternating gray and white shading, with the center  $T_K = 0$  of each corresponding superluminal kink shown with a superimposed curve. Left:  $\epsilon_N = 10^{-5}$ . Right:  $\epsilon_N = 10^{-7}$ . In both cases  $4\nu^{1/3} = 1$ . For different values of  $\nu > 0$ , the curves can “march out” of the confining strips as  $|t|$  increases, unless  $\epsilon_N$  is small enough.

We may now formulate some observations about the kink centered at  $T_K = 0$  that follow from Proposition 1.1. Indeed, for each  $m \in \mathbb{Z}$  the following is true: for  $z$  sufficiently negative,  $\sigma := \text{sgn}(\mathcal{U}_m(z)) = 1$ , while for  $z$  sufficiently positive,  $\sigma = (-1)^m$ . This then implies that as  $z \rightarrow -\infty$  we have a series of time-like antikinks (all of the same topological charge) consistent with nearly synchronous rotational motion of pendula. Similarly, as  $z \rightarrow +\infty$  we have alternation between kink and antikink with no net topological charge consistent with nearly synchronous librational motion of pendula. These facts demonstrate that the asymptotic behavior near criticality matches appropriately with the established asymptotic formulae [7] valid away from criticality that model  $u_N(x, t)$  as a modulated train of superluminal rotational traveling waves to the left of  $x_{\text{crit}}$  and as a modulated train of superluminal librational traveling waves to the right of  $x_{\text{crit}}$ . It also follows from Proposition 1.1 that if  $m > 0$  then  $\mathcal{U}_m(z)$  blows up as  $z \rightarrow \pm\infty$  while if  $m < 0$  then  $\mathcal{U}_m(z) \rightarrow 0$  as  $z \rightarrow \pm\infty$ . This implies that for  $m > 0$  the corresponding kinks follow a logarithmic “frown” for large  $|z|$ , while for  $m < 0$  they instead follow a logarithmic “smile” for large  $|z|$ . These features can be seen in Figure 1.8.

When the horizontal strips of the  $(x, t)$ -plane corresponding to different integral values of  $m$  are put together, one sees that the kink-like asymptotics given by Theorem 1.3 are valid throughout the region where  $\Delta x = \mathcal{O}(\epsilon_N^{2/3})$  and  $t = \mathcal{O}(\epsilon_N \log(\epsilon_N^{-1}))$  with the sole exception of small sub-regions near the top and bottom edges of each strip where the curve  $T_K = 0$  tries to exit the strip. When the strips are put together there are obvious mismatches of the curves from neighboring strips in these small sub-regions as can be seen



from the left-hand plot in Figure 1.8. Our final result corrects the kink asymptotics in these regions and therefore removes the mismatches.

**Theorem 1.4** (Grazing kink collisional asymptotics). *Fix an integer  $m$  and let  $z_0$  denote any of the (necessarily simple) real zeros of  $\mathcal{U}_{m-1}(z)$  (equivalently a simple zero of  $\mathcal{V}_{m+1}(z)$ ). Given a sufficiently small positive number  $\mu > 0$ , suppose that  $|t - (\frac{2}{3}m - \frac{1}{3})\epsilon_N \log(\epsilon_N^{-1})| \leq \frac{1}{3}\epsilon_N \log(\epsilon_N^{-1})$ , and that  $|z - z_0| \leq \mu\epsilon_N^{1/6} \exp(|t - (\frac{2}{3}m - \frac{1}{3})\epsilon_N \log(\epsilon_N^{-1})|/(2\epsilon_N))$ , where  $z$  is defined by (1.43). Then*

$$\begin{aligned} \cos(\tfrac{1}{2}u_N(x, t)) &= (-1)^{m-1} \operatorname{sgn}(\mathcal{U}'_{m-1}(z_0)) \frac{2X_G \operatorname{sech}(T_G)}{1 + X_G^2 \operatorname{sech}^2(T_G)} + E_{\cos}(x, t; \epsilon_N) \\ \sin(\tfrac{1}{2}u_N(x, t)) &= (-1)^{m-1} \frac{1 - X_G^2 \operatorname{sech}^2(T_G)}{1 + X_G^2 \operatorname{sech}^2(T_G)} + E_{\sin}(x, t; \epsilon_N), \end{aligned} \quad (1.48)$$

where

$$X_G := 2 \left( \frac{\nu}{\epsilon_N} \right)^{1/3} \left( \frac{\Delta x}{2\nu^{1/3}\epsilon_N^{2/3}} - z_0 \right) \quad \text{and} \quad T_G := \frac{t}{\epsilon_N} - (2m-1) \log \left( \frac{4\nu^{1/3}}{\epsilon_N^{1/3}} \right) + \log |\mathcal{U}'_{m-1}(z_0)|. \quad (1.49)$$

The error terms satisfy

$$\lim_{\epsilon_N \downarrow 0} E_{\cos}(x, t; \epsilon_N) = \lim_{\epsilon_N \downarrow 0} E_{\sin}(x, t; \epsilon_N) = 0 \quad (1.50)$$

whenever  $|z - z_0| \ll 1$  as  $\epsilon_N \downarrow 0$  (which excludes only the four extreme corners of the hourglass-shaped region under consideration). Moreover, we have

$$E_{\cos}(x, t; \epsilon_N) = \mathcal{O}(\epsilon_N^{1/3}) \quad \text{and} \quad E_{\sin}(x, t; \epsilon_N) = \mathcal{O}(\epsilon_N^{1/3}) \quad (1.51)$$

holding uniformly for  $|z - z_0| = \mathcal{O}(\epsilon_N^{1/3})$  and  $|t - (\frac{2}{3}m - \frac{1}{3})\epsilon_N \log(\epsilon_N^{-1})| \leq \frac{1}{3}\epsilon_N \log(\epsilon_N^{-1})$ .

Wherever the error terms vanish in the limit, it follows from (1.48) that  $\cos(u_N(x, t)) \rightarrow \cos(u(X_G, T_G))$  and  $\sin(u_N(x, t)) \rightarrow \sin(u(X_G, T_G))$ , where with  $\kappa := \operatorname{sgn}(\mathcal{U}'_{m-1}(z_0))$ , a function  $u(X, T)$  is defined modulo  $2\pi$  by

$$\begin{aligned} \cos(u(X, T)) &:= \frac{8X^2 \operatorname{sech}^2(T)}{(1 + X^2 \operatorname{sech}^2(T))^2} - 1 \\ \sin(u(X, T)) &:= 4\kappa X \operatorname{sech}(T) \frac{1 - X^2 \operatorname{sech}^2(T)}{(1 + X^2 \operatorname{sech}^2(T))^2}. \end{aligned} \quad (1.52)$$

Then it is again easy to confirm that  $u(X, T)$  is an exact solution of the unscaled sine-Gordon equation (1.47). Plots of  $\cos(u(X, T))$  and  $\sin(u(X, T))$  are shown in Figure 1.9. This is a particular solution of the sine-Gordon equation that corresponds to boundary conditions of the form  $u \rightarrow \pi \pmod{2\pi}$  as  $T \rightarrow \pm\infty$ . In the proper version of scattering theory corresponding to these boundary conditions, the solution at hand is a reflectionless potential associated to a *double eigenvalue*. In the context of the Riemann-Hilbert problem of inverse scattering such an object is encoded as a double pole of the matrix unknown. Such solutions were noted in the earliest days of inverse scattering by Zakharov and Shabat [24] in their study of the focusing nonlinear Schrödinger (NLS) equation. They pointed out that such solutions describe *grazing collisions* of solitons. The trajectories of the solitons emerging from the interaction region are not asymptotically straight lines, but rather are logarithmic, with relative velocities of the solitons tending to zero. In the context of the sine-Gordon equation we have instead a grazing collision of superluminal kinks. These structures serve to smooth out the mismatches of curves shown in Figure 1.8 at horizontal strip boundaries. In the language of matched asymptotic expansions, they function as internal transition layers of “corner” type.

The regions of validity of the asymptotic formulae given in Theorems 1.3 and 1.4 are compared in Figure 1.10, which illustrates the complementary nature of these two results. It is clear from this figure that Theorems 1.3 and 1.4 both provide simple asymptotic formulae for the same fields in overlapping regions. In the overlap domains, formulae of kink type provided by Theorem 1.3, and of grazing kink collision type provided by Theorem 1.4 are *simultaneously valid*. This is actually a consequence of Theorem 1.2, but it may also be checked directly, by writing both types of formulae in terms of common spatiotemporal independent variables.

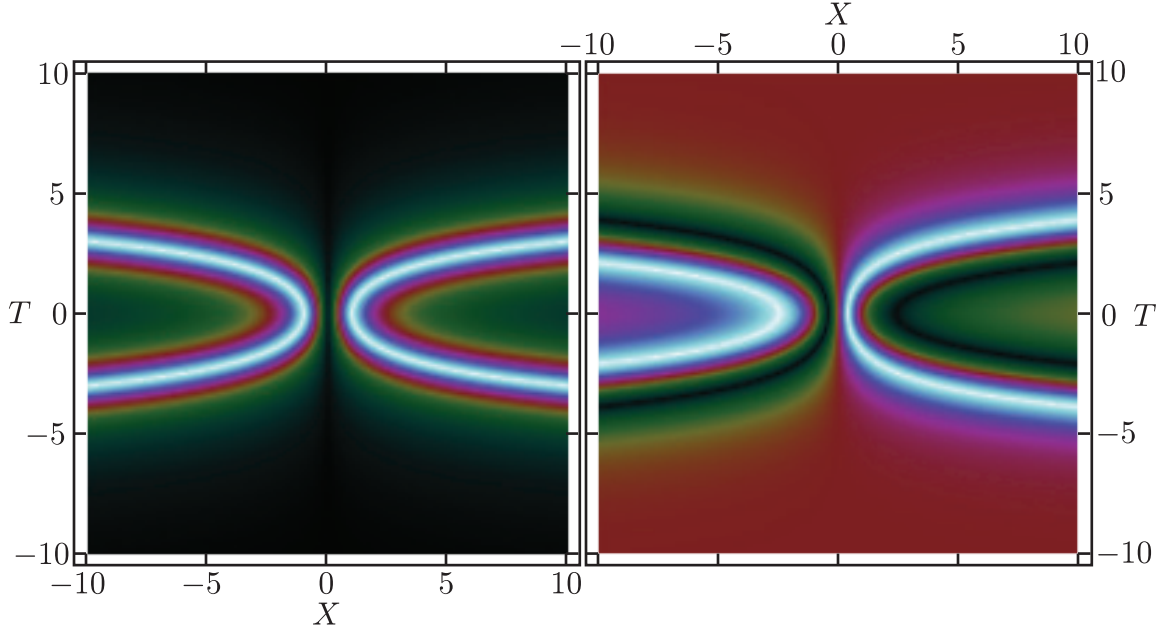


FIGURE 1.9. (*Grazing collision of superluminal kinks*). The cosine (left) and sine (right) of  $u(X, T)$  defined by (1.52) with  $\kappa = 1$ . For  $\kappa = -1$  the plots are simply reflected through the vertical  $T$ -axis.

One way to characterize the asymptotic formulae giving the universal form of the dispersive breakup of the simple pendulum separatrix under the sine-Gordon equation is to say that the wave pattern consists of waves of elementary “solitonic” forms that are spatiotemporally arranged according to solutions of certain nonlinear differential equations of Painlevé type. To our knowledge, this type of phenomenon was first observed quite recently in a paper by Bertola and Tovbis [5] in which all of the local maxima of the modulus of the solution to the focusing NLS equation in the semiclassical limit near the onset of oscillatory behavior (the point of so-called elliptic umbilic catastrophe of the approximating elliptic quasilinear Whitham modulational system) are individually modeled by the same exact solution of the focusing NLS equation (the rational or Peregrine breather solution) but the locations of the maxima are far more transcendental, being determined by the location of poles in the complex plane of the so-called *tritronquée* solution of the Painlevé-I equation. In the current context, the role of the tritronquée solution is played by the rational functions  $\{\mathcal{U}_m\}_{m \in \mathbb{Z}}$  solving the Painlevé-II system, and interestingly, *we require an infinite number of different Painlevé functions to describe the full wave pattern*.

## 2. CHOICE OF $g(w)$

The use of a so-called “ $g$ -function” to precondition a matrix Riemann-Hilbert problem for subsequent asymptotic analysis by the Deift-Zhou steepest-descent method is by now a standard tool, having first been introduced by Deift, Venakides, and Zhou [15] in their analysis of the Korteweg-de Vries (KdV) equation in the small-dispersion limit. Recall that the scalar function  $g$  appearing in the formulation of Riemann-Hilbert Problem 1.1 is subject to the basic requirements of

- Analyticity:  $g$  is analytic in the domain  $\mathbb{C} \setminus (P_\infty \cup \mathbb{R}_+)$ ,
- Boundary behavior:  $g$  takes well-defined boundary values on  $P_\infty \cup \mathbb{R}_+$  in the classical sense, and the sum of the boundary values vanishes on  $\mathbb{R}_+$ , and,
- Normalization:  $g(0) = g(\infty) = 0$ .

We can use any such function to formulate the basic Riemann-Hilbert problem for  $\mathbf{N}(w) = \mathbf{N}_N^g(w; x, t)$ , but it is well-known that to study the limit  $N \rightarrow \infty$  it is important that  $g$  be chosen appropriately. In [7] it is explained how  $g$  should be chosen given  $(x, t) \in \mathbb{R}^2$  to calculate the limit  $N \rightarrow \infty$  with  $(x, t)$

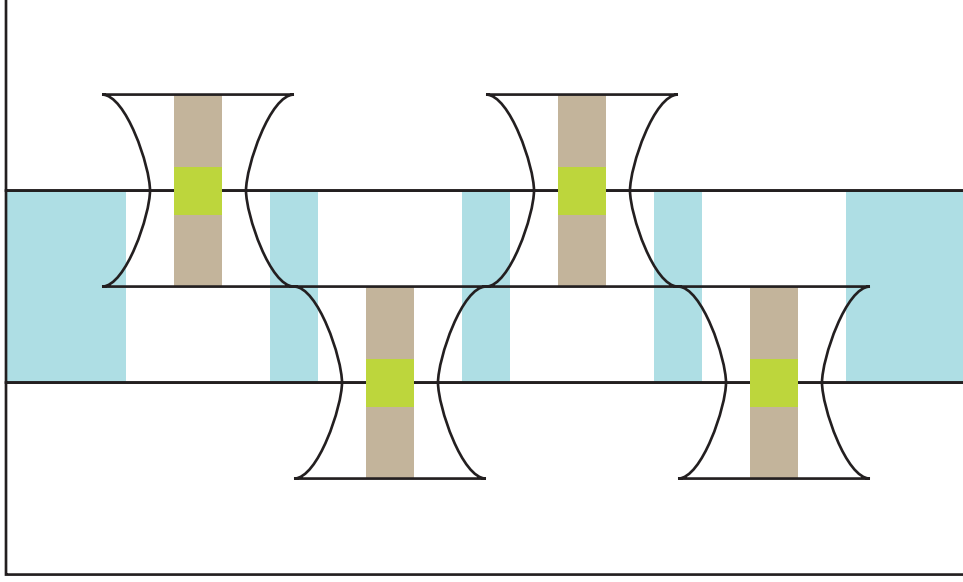


FIGURE 1.10. *Theorem 1.3 provides asymptotics in horizontal strips in the  $(z, t)$ -plane with the exception of the small rectangles shown in green, and with optimal accuracy of  $\mathcal{O}(\epsilon_N^{1/3})$  in the subregions shown in blue. Theorem 1.4 provides asymptotics in hourglass-shaped regions with the exception of where they overlap the blue regions (where the error estimate of Theorem 1.3 is optimally small) and with optimal accuracy of  $\mathcal{O}(\epsilon_N^{1/3})$  in the included green regions (where the error estimate of Theorem 1.3 fails to be small) and brown regions.*

fixed. The chosen function  $g(w) = g(w; x, t)$  has the property that its boundary values satisfy a kind of equilibrium condition (an analogue of (2.2) below) on a contour (called  $\beta$  in [7]) whose topology changes at criticality: in the simplest (genus one) case, near  $(x, t) = (x_{\text{crit}}, 0)$  the equilibrium contour is either a nearly circular arc through  $w = 1$  with complex-conjugate endpoints near  $w = -1$  (this corresponds to modulated superluminal librational waves) or the union of a nearly circular closed contour with a transecting small interval of the real axis with both endpoints near  $w = -1$  (this corresponds to modulated superluminal rotational waves). Exactly at criticality the two endpoints coalesce at  $w = -1$  and the equilibrium contour becomes the unit circle. It turns out that proximity of endpoints of this contour is an obstruction to proving uniform asymptotics for  $(x, t)$  near, but not exactly at, criticality.

Here, since we want to allow  $(x, t)$  to approach the critical point  $(x_{\text{crit}}, 0)$  at some rate depending on  $N$ , we need to use a different function than  $g(w) = g(w; x, t)$  defined in [7]. We will replace  $g(w)$  with  $\mathbf{g}(w) = \mathbf{g}(w; x, t)$  whose construction and properties are described below. It is important that while for general  $x$  and  $t$ , the functions  $g(w)$  and  $\mathbf{g}(w)$  are different, at criticality they coincide (they satisfy the same conditions, which can be shown to uniquely determine the solution).

The function  $\mathbf{g}(w)$  is required to satisfy the following conditions:

- $\mathbf{g}(w)$  is analytic if  $|w| \neq 1$  and  $w \notin \mathbb{R}_+$ , and it takes continuous and bounded boundary values on the unit circle  $|w| = 1$  and on the positive half-line  $\mathbb{R}_+$ .
- $\mathbf{g}(w)$  obeys the following jump conditions:

$$\mathbf{g}_+(\xi) + \mathbf{g}_-(\xi) = 0, \quad \xi \in \mathbb{R}_+. \quad (2.1)$$

$$2iQ(\xi; x, t) + \bar{L}(\xi) - \mathbf{g}_+(\xi) - \mathbf{g}_-(\xi) = 0, \quad |\xi| = 1. \quad (2.2)$$

- $\mathbf{g}(w)$  has the following values:

$$\mathbf{g}(0) = 0. \quad (2.3)$$

$$\lim_{w \rightarrow \infty} \mathbf{g}(w) = 0. \quad (2.4)$$

To obtain a formula for  $\mathbf{g}(w)$ , we first write  $\mathbf{g}$  in the form

$$\mathbf{g}(w) = \frac{1}{2}L(w) + \mathfrak{h}(w), \quad (2.5)$$

where the function  $L$  is defined by (1.15). Therefore the above conditions on  $\mathbf{g}$  are translated into equivalent conditions on  $\mathfrak{h}$ :

- $\mathfrak{h}(w)$  is analytic if  $|w| \neq 1$ ,  $w \notin \mathbb{R}_+$ , and  $w \notin [\mathfrak{a}, \mathfrak{b}]$ , and it takes continuous boundary values on the unit circle  $|w| = 1$ , on the positive half-line  $\mathbb{R}_+$ , and on the interval  $[\mathfrak{a}, \mathfrak{b}]$ .
- $\mathfrak{h}(w)$  obeys the following jump conditions:

$$\mathfrak{h}_+(\xi) + \mathfrak{h}_-(\xi) = 0, \quad \xi \in \mathbb{R}_+. \quad (2.6)$$

$$2iQ(\xi; x, t) - \mathfrak{h}_+(\xi) - \mathfrak{h}_-(\xi) = 0, \quad |\xi| = 1. \quad (2.7)$$

$$\mathfrak{h}_+(\xi) - \mathfrak{h}_-(\xi) = -\frac{1}{2}(L_+(\xi) - L_-(\xi)) = -i\theta_0(\xi), \quad w \in [\mathfrak{a}, \mathfrak{b}]. \quad (2.8)$$

In the latter condition, the two complementary segments  $(\mathfrak{a}, -1)$  and  $(-1, \mathfrak{b})$  are taken to be oppositely oriented toward  $\xi = -1$ .

- $\mathfrak{h}(w)$  has the following values:

$$\mathfrak{h}(0) = 0. \quad (2.9)$$

$$\lim_{w \rightarrow \infty} \mathfrak{h}(w) = 0. \quad (2.10)$$

Finally, we write  $\mathfrak{h}(w)$  in terms of a new equivalent unknown  $\mathbf{m}(w)$  by the piecewise substitution:

$$\mathfrak{h}(w) := \begin{cases} (-w)^{-1/2}\mathbf{m}(w), & |w| > 1 \\ -(-w)^{-1/2}\mathbf{m}(w), & |w| < 1. \end{cases} \quad (2.11)$$

The conditions satisfied by  $\mathbf{m}(w)$  are then the following:

- $\mathbf{m}(w)$  is analytic if  $|w| \neq 1$  and  $w \notin [\mathfrak{a}, \mathfrak{b}]$ , and it takes continuous boundary values on the unit circle  $|w| = 1$  and the interval  $[\mathfrak{a}, \mathfrak{b}]$ .
- $\mathbf{m}(w)$  obeys the following jump conditions:

$$\mathbf{m}_+(\xi) - \mathbf{m}_-(\xi) = -2i(-\xi)^{1/2}Q(\xi; x, t), \quad |\xi| = 1, \quad (2.12)$$

where the orientation of the circle is in the positive (counterclockwise) sense, and

$$\mathbf{m}_+(\xi) - \mathbf{m}_-(\xi) = -i(-\xi)^{1/2}\theta_0(\xi), \quad \xi \in [\mathfrak{a}, \mathfrak{b}], \quad (2.13)$$

where now the interval  $[\mathfrak{a}, \mathfrak{b}]$  is taken to be oriented left-to-right. Note that  $(-\xi)^{1/2} > 0$  for  $\mathfrak{a} < \xi < \mathfrak{b} < 0$ , and that the right-hand side of (2.12) is a single-valued function on the circle:

$$-2i(-\xi)^{1/2}Q(\xi; x, t) = \frac{x}{2}(1 - \xi) - \frac{t}{2}(1 + \xi), \quad |\xi| = 1. \quad (2.14)$$

- $\mathbf{m}(w)$  is uniformly bounded, and has the following value:

$$\mathbf{m}(0) = 0. \quad (2.15)$$

We may now express  $\mathbf{m}(w)$  in terms of a Cauchy integral using the Plemelj formula:

$$\mathbf{m}(w) = \mathbf{m}_0 + \frac{1}{2\pi i} \oint_{|\xi|=1} \frac{-2i(-\xi)^{1/2}Q(\xi; x, t)}{\xi - w} d\xi + \frac{1}{2\pi i} \int_{\mathfrak{a}}^{\mathfrak{b}} \frac{-i(-\xi)^{1/2}\theta_0(\xi)}{\xi - w} d\xi, \quad (2.16)$$

where the constant  $\mathbf{m}_0$  is the most general entire function we can add to the Cauchy integrals consistent with the uniform boundedness of  $\mathbf{m}(w)$ . The integral over the positively-oriented unit circle may be evaluated in closed form:

$$\mathbf{m}(w) = \mathbf{m}_0 + \frac{1}{2\pi i} \int_{\mathfrak{a}}^{\mathfrak{b}} \frac{-i(-\xi)^{1/2}\theta_0(\xi)}{\xi - w} d\xi + \begin{cases} \frac{x}{2}(1 - w) - \frac{t}{2}(1 + w), & |w| < 1 \\ 0, & |w| > 1. \end{cases} \quad (2.17)$$

We choose the value of the constant  $\mathbf{m}_0$  to enforce the condition (2.15):

$$\mathbf{m}_0 = \frac{t - x}{2} - \frac{1}{2\pi} \int_{\mathfrak{a}}^{\mathfrak{b}} \frac{\theta_0(\xi)}{(-\xi)^{1/2}} d\xi. \quad (2.18)$$

It follows that

$$\mathbf{m}(w) = \frac{w}{2\pi} \int_a^b \frac{\theta_0(\xi) d\xi}{(-\xi)^{1/2}(\xi - w)} + \begin{cases} -\frac{x+t}{2}w, & |w| < 1 \\ \frac{t-x}{2}, & |w| > 1. \end{cases} \quad (2.19)$$

By writing  $\mathbf{g}$  in terms of  $\mathbf{h}$  (by (2.5)) and then writing  $\mathbf{h}$  in terms of  $\mathbf{m}$  (by (2.11)), this completes the construction of  $\mathbf{g}(w)$ .

### 3. STEEPEST DESCENT, THE OUTER MODEL PROBLEM, AND ITS SOLUTION

According to (2.2), with  $\mathbf{N}(w) = \mathbf{N}_N^{\mathbf{g}}(w; x, t)$  and with  $\mathbf{g}(w) = \mathbf{g}(w; x, t)$  defined as in §2, a condition of “equilibrium type” holds on the unit circle  $|w| = 1$ , regardless of the values of  $x$  and  $t$ . The next step is to “open a lens” about the whole circle. The lens consists of four disjoint open sets separated by the unit circle and the interval  $I \subset \mathbb{R}_+$  as shown with shading in Figure 3.1. The union of the two components lying on the left (respectively right) of the unit circle with its orientation will be denoted  $\Lambda^+$  (respectively  $\Lambda^-$ ). We now define a new unknown  $\mathbf{O}(w)$  equivalent to  $\mathbf{N}(w)$  by the substitution

$$\mathbf{O}(w) := \begin{cases} \mathbf{N}(w)T_N(w)^{-\sigma_3/2} \begin{bmatrix} 1 & \mp i e^{-[2iQ(w; x, t) + L(w) \mp i\theta_0(w) - 2\mathbf{g}(w)]/\epsilon_N} \\ 0 & 1 \end{bmatrix}, & w \in \Lambda^\pm \\ \mathbf{N}(w), & w \in \mathbb{C} \setminus (\Sigma_{\mathbf{N}} \cup \Lambda^+ \cup \Lambda^-). \end{cases} \quad (3.1)$$

Here, the square root  $T_N(w)^{1/2}$  is well-defined as the principal branch, since the uniform approximation  $T_N(w) \approx 1$  holds when  $N$  is large. (The definition (3.1) coincides with that of  $\mathbf{O}(w)$  in terms of  $\mathbf{N}(w)$  given in [7] in the case that  $\Delta = \emptyset$  and with  $g$  replaced by  $\mathbf{g}$ .) The jump contour for  $\mathbf{O}(w)$  is the same as that for  $\mathbf{N}(w)$  but augmented by four arcs emanating from  $w = -1$  representing the “outer” boundaries of the lens halves  $\Lambda^\pm$  as shown in Figure 3.1.

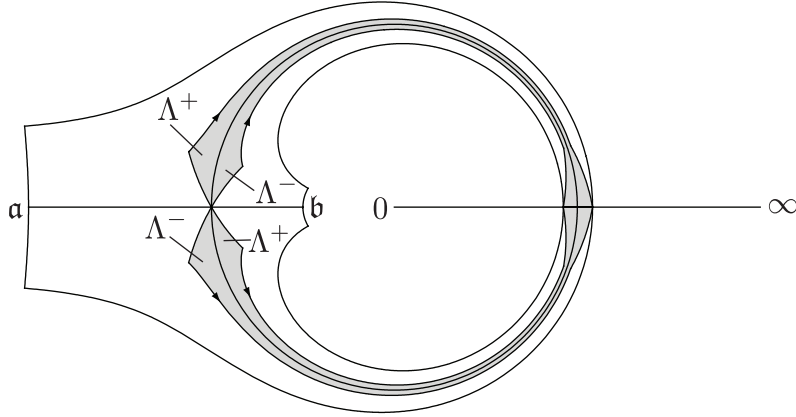


FIGURE 3.1. The contour of discontinuity for the sectionally analytic function  $\mathbf{O}(w)$ , with the lens halves  $\Lambda^\pm$  indicated with shading. Arcs of the jump contour also present in the jump contour for  $\mathbf{N}(w)$  (see Figure 1.6) retain that orientation, and the lens boundaries are oriented as indicated.

It is a consequence of the equilibrium condition (2.2) satisfied by the boundary values taken by  $\mathbf{g}(w)$  on the unit circle that  $\mathbf{O}(w)$  satisfies the piecewise-constant jump conditions

$$\mathbf{O}_+(\xi) = \mathbf{O}_-(\xi)(\mp i\sigma_1), \quad |\xi| = 1, \quad \pm \Im\{\xi\} > 0. \quad (3.2)$$



Since the equilibrium contour is a closed curve, we may remove these discontinuities from the problem by making another explicit transformation:

$$\mathbf{P}(w) := \begin{cases} \mathbf{O}(w), & |w| > 1 \\ \mathbf{O}(w)(-i\sigma_1), & |w| < 1 \text{ and } \Im\{w\} > 0 \\ \mathbf{O}(w)i\sigma_1, & |w| < 1 \text{ and } \Im\{w\} < 0. \end{cases} \quad (3.3)$$

Unlike  $\mathbf{O}(w)$ , the matrix function  $\mathbf{P}(w)$  extends continuously to the unit circle, but the jump conditions it satisfies within the unit disk where it differs from  $\mathbf{O}(w)$  are altered somewhat from those of  $\mathbf{O}(w)$ , including a new jump of the simple form  $\mathbf{P}_+(\xi) = -\mathbf{P}_-(\xi)$  on the segment  $\mathfrak{b} < \xi < 0$ . The contour of discontinuity for  $\mathbf{P}(w)$  is illustrated in Figure 3.2.

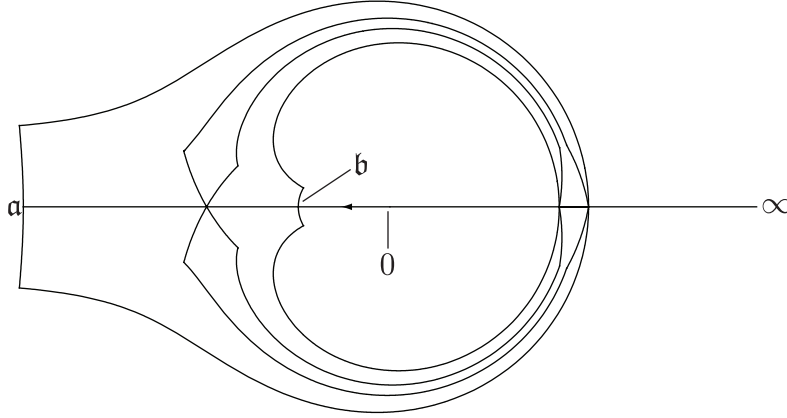


FIGURE 3.2. The contour of discontinuity for the sectionally analytic function  $\mathbf{P}(w)$ . Contour arcs also contained in the jump contour for  $\mathbf{O}(w)$  (see Figures 1.6 and 3.1) retain that orientation, and otherwise (that is, in the interval  $(\mathfrak{b}, 0)$ ) the orientation is as indicated.

The jump discontinuities of  $\mathbf{P}(w)$  will turn out to be negligible except in a small neighborhood of  $w = -1$  and along the ray  $w > -1$ , and  $\mathbf{P}(w)$  is a matrix tending to the identity as  $w \rightarrow \infty$ . These considerations lead us to propose a model Riemann-Hilbert problem whose solution we expect to approximate  $\mathbf{P}(w)$  away from  $w = -1$ :

**Riemann-Hilbert Problem 3.1** (Outer model problem near criticality). *Let  $w_* \approx -1$  be a real parameter. Find a  $2 \times 2$  matrix  $\dot{\mathbf{P}}^{\text{out}}(w)$  with the following properties.*

**Analyticity:**  $\dot{\mathbf{P}}^{\text{out}}(w)$  is an analytic function of  $w$  for  $w \in \mathbb{C} \setminus [w_*, +\infty)$ , Hölder continuous up to the jump interval  $[w_*, +\infty)$  with the exception of an arbitrarily small neighborhood of the point  $w = w_*$ . In a neighborhood of  $w = w_*$ , the elements of  $\dot{\mathbf{P}}^{\text{out}}(w)$  are bounded by an unspecified power of  $|w - w_*|$ .

**Jump condition:** The boundary values taken by  $\dot{\mathbf{P}}^{\text{out}}(w)$  along  $(w_*, 0)$  and  $\mathbb{R}_+$  satisfy the following jump conditions:

$$\dot{\mathbf{P}}_+^{\text{out}}(\xi) = -\dot{\mathbf{P}}_-^{\text{out}}(\xi), \quad w_* < \xi < 0, \quad (3.4)$$

and

$$\dot{\mathbf{P}}_+^{\text{out}}(\xi) = \sigma_2 \dot{\mathbf{P}}_-^{\text{out}}(\xi) \sigma_2, \quad \xi \in \mathbb{R}_+. \quad (3.5)$$

As these jump conditions are both involutive, orientation of the jump contours is irrelevant.

**Normalization:** The following normalization conditions hold:

$$\lim_{w \rightarrow \infty} \dot{\mathbf{P}}^{\text{out}}(w) = \mathbb{I} \quad \text{and} \quad \det(\dot{\mathbf{P}}^{\text{out}}(w)) \equiv 1. \quad (3.6)$$

There are many solutions of this Riemann-Hilbert problem, a fact that can be traced partly to the unspecified power-law rate of growth admitted as  $w \rightarrow w_*$ . The complete variety of solutions is not directly relevant for us; we will simply select a family of particular solutions in the class of diagonal matrices. Indeed, for each  $m \in \mathbb{Z}$ , we have the solution

$$\dot{\mathbf{P}}^{\text{out}}(w) = \dot{\mathbf{P}}_m^{\text{out}}(w) := \left[ (-w)^{1/2} + (-w_*)^{1/2} \right]^{(1-2m)\sigma_3} (w_* - w)^{(m-1/2)\sigma_3}. \quad (3.7)$$

Note that if  $m$  is held fixed, then  $\dot{\mathbf{P}}_m^{\text{out}}(w)$  and its inverse are bounded when  $w$  is bounded away from  $w_*$ . Later (see (4.52)), we will let  $w_*$  depend weakly on  $x$  and  $t$  in such a way that  $w_* \rightarrow -1$  as  $t \rightarrow 0$  and  $x \rightarrow x_{\text{crit}}$ . Clearly, any bound for  $\dot{\mathbf{P}}_m^{\text{out}}(w)$  or its inverse valid for  $w$  bounded away from  $-1$  will hold uniformly with respect to  $x$  and  $t$  near criticality, and of course  $\dot{\mathbf{P}}_m^{\text{out}}(w)$  is independent of  $\epsilon_N$ .

#### 4. INNER MODEL PROBLEM VALID NEAR $w = -1$

**4.1. Exact jump matrices for  $\mathbf{P}(w)$  near  $w = -1$ .** Let  $U$  be a fixed neighborhood of  $w = -1$ . By straightforward substitutions, we may assume without loss of generality that within  $U$ , the jump contour for  $\mathbf{P}(w)$  consists of the real axis together with four arcs (the lens boundaries) meeting at some real point  $w = w_* \in U \cap \mathbb{R}$  tending to  $-1$  as criticality is approached; the way  $w_*$  is determined as a function of  $x$  and  $t$  will be explained later (the resulting formula being (4.52)). We suppose further that the lens boundaries lie along certain curves (also to be specified later) tangent at  $w = w_*$  to the straight lines  $\arg(w - w_*) = \pm\pi/3$  and  $\arg(w - w_*) = \pm 2\pi/3$ . Some calculations show that the exact jump conditions for  $\mathbf{P}(w)$  within  $U$  can be expressed only in terms of the analytic function  $T_N(w) \approx 1$  and another analytic function  $k(w)$ , which is the analytic continuation from  $w < \min\{w_*, -1\}$  of the function

$$k(w) = k(w; x, t) := 2iQ(w; x, t) + \bar{L}(w) - 2\mathfrak{g}(w; x, t), \quad w < -1. \quad (4.1)$$

The jump matrix  $\mathbf{V}_{\mathbf{P}}(\xi)$  for which along each of six contour arcs meeting at  $w = w_*$  we have  $\mathbf{P}_+(\xi) = \mathbf{P}_-(\xi)\mathbf{V}_{\mathbf{P}}(\xi)$  is illustrated in Figure 4.1. The function  $T_N(w)$  can easily be removed from the jump conditions

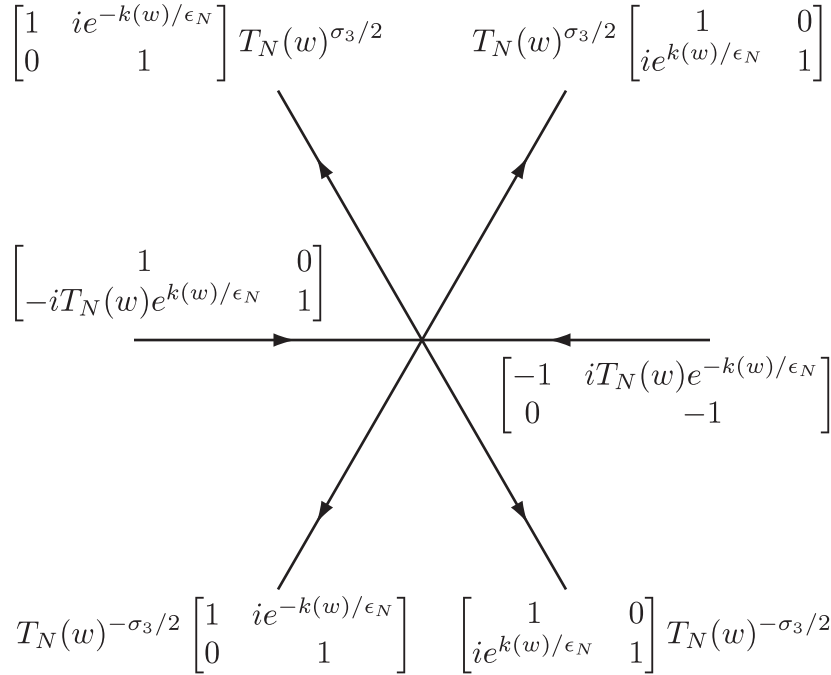


FIGURE 4.1. The jump matrix for  $\mathbf{P}(w)$  near  $w = w_* \approx -1$ . The four non-real arcs of the jump contour are in general only approximately linear near the intersection point  $w = w_*$ .

by making the following near-identity transformation (locally, in  $U$ ):

$$\tilde{\mathbf{P}}(w) := \mathbf{P}(w)d_N(w)^{\sigma_3} \quad (4.2)$$

where  $d_N(w)$  is the piecewise-analytic function given by

$$d_N(w) := \begin{cases} T_N(w)^{-1/2}, & \text{in the region tangent to the sector } |\arg(w_* - w)| < \frac{\pi}{3} \\ T_N(w)^{1/2}, & \text{in the region tangent to the sector } |\arg(w - w_*)| < \frac{\pi}{3} \\ 1, & \text{otherwise.} \end{cases} \quad (4.3)$$

Note that  $d_N(w) = 1 + \mathcal{O}(\epsilon_N)$  holds uniformly for  $w \in U$ , so indeed  $\mathbf{P}(w)^{-1}\tilde{\mathbf{P}}(w) = \mathbb{I} + \mathcal{O}(\epsilon_N)$ . The jump conditions satisfied by  $\tilde{\mathbf{P}}(w)$  near  $w = w_* \approx -1$  are as shown in Figure 4.1 but with  $T_N(w)$  replaced by 1 in all cases.

#### 4.2. Expansion of $k(w)$ about $w = -1$ .

**Proposition 4.1.** *The function  $k(w) = k(w; x, t)$  is analytic at  $w = -1$ , having the Taylor expansion*

$$\begin{aligned} k(w; x, t) = & -t + \frac{1}{2}\Delta x(w+1) + \left[\frac{1}{4}\Delta x - \frac{1}{8}t\right](w+1)^2 \\ & + \left[\frac{3}{16}\Delta x - \frac{1}{8}t + \nu\right](w+1)^3 \\ & + \left[\frac{5}{32}\Delta x - \frac{15}{128}t + \frac{3}{2}\nu\right](w+1)^4 + \mathcal{O}((w+1)^5), \quad w \rightarrow -1, \end{aligned} \quad (4.4)$$

where  $\nu > 0$  is independent of  $w$ ,  $x$ , and  $t$ , being defined by (1.30). All Taylor coefficients are linear in  $\Delta x$  and  $t$ , and hence in particular the error term is uniform for bounded  $\Delta x$  and  $t$ .

*Proof.* For  $w < \min\{w_*, -1\}$  and  $w \in U$ , we have the formula

$$k(w; x, t) = 2iQ(w; x, t) - \frac{1}{(-w)^{1/2}}[\mathbf{m}_+(w) + \mathbf{m}_-(w)], \quad (4.5)$$

where  $\mathbf{m}(w)$  is given by

$$\mathbf{m}(w) = \frac{t-x}{2} + \frac{w}{2\pi} \int_{\mathbf{a}}^{\mathbf{b}} \frac{\theta_0(s) ds}{(-s)^{1/2}(s-w)}. \quad (4.6)$$

To express the boundary values  $\mathbf{m}_{\pm}(w)$  for  $w < \min\{w_*, -1\}$  in a form admitting analytic continuation to a full neighborhood of  $w = -1$  (a point that lies on the discontinuity contour  $[\mathbf{a}, \mathbf{b}]$  of  $\mathbf{m}(w)$ ), let  $C_{\pm}$  denote two contours from  $s = \mathbf{a}$  to  $s = \mathbf{b}$ , with  $C_+$  in the upper half  $s$ -plane and  $C_-$  in the lower half  $s$ -plane, and such that the two contours are mapped onto each other with reversal of orientation under the mapping  $s \mapsto s^{-1}$  (in particular, this mapping permutes the points  $\mathbf{a}$  and  $\mathbf{b}$ ). Then, also recalling the definition of  $Q(w; x, t)$  in terms of  $E(w)$  and  $D(w)$ ,

$$k(w) = -\frac{x}{2} \left( (-w)^{1/2} - (-w)^{-1/2} \right) - \frac{t}{2} \left( (-w)^{1/2} + (-w)^{-1/2} \right) + \frac{(-w)^{1/2}}{2\pi} \int_{C_+ \cup C_-} \frac{\theta_0(s) ds}{(-s)^{1/2}(s-w)}. \quad (4.7)$$

This formula represents the analytic continuation to a full neighborhood of  $w = -1$  of the function  $k$  originally defined by (4.1) only for  $w < \min\{w_*, -1\}$ . In fact, the domain of analyticity for  $k$  has now been extended to the domain enclosed by the contour  $C_+ \cup C_-$ . It is also obvious that  $k$  depends linearly on  $x$  and  $t$ , and this property will clearly be inherited by all of its Taylor coefficients.

To analyze  $k(w)$  near the point  $w = -1$  to which the self-intersection point  $w_*$  will converge at criticality, we begin with the following elementary Taylor expansions (convergent for  $|w+1| < 1$ ):

$$\begin{aligned} -\frac{1}{2} \left( (-w)^{1/2} - (-w)^{-1/2} \right) &= \frac{1}{2}(w+1) + \frac{1}{4}(w+1)^2 + \frac{3}{16}(w+1)^3 + \frac{5}{32}(w+1)^4 + \mathcal{O}((w+1)^5) \\ -\frac{1}{2} \left( (-w)^{1/2} + (-w)^{-1/2} \right) &= -1 - \frac{1}{8}(w+1)^2 - \frac{1}{8}(w+1)^3 - \frac{15}{128}(w+1)^4 + \mathcal{O}((w+1)^5). \end{aligned} \quad (4.8)$$

We also have the expansion

$$\begin{aligned}
\frac{(-w)^{1/2}}{s-w} &= \frac{1}{s+1} + \left[ \frac{1}{(s+1)^2} - \frac{1}{2} \frac{1}{s+1} \right] (w+1) \\
&+ \left[ \frac{1}{(s+1)^3} - \frac{1}{2} \frac{1}{(s+1)^2} - \frac{1}{8} \frac{1}{s+1} \right] (w+1)^2 \\
&+ \left[ \frac{1}{(s+1)^4} - \frac{1}{2} \frac{1}{(s+1)^3} - \frac{1}{8} \frac{1}{(s+1)^2} - \frac{1}{16} \frac{1}{s+1} \right] (w+1)^3 \\
&+ \left[ \frac{1}{(s+1)^5} - \frac{1}{2} \frac{1}{(s+1)^4} - \frac{1}{8} \frac{1}{(s+1)^3} - \frac{1}{16} \frac{1}{(s+1)^2} - \frac{5}{128} \frac{1}{s+1} \right] (w+1)^4 \\
&+ \mathcal{O}((w+1)^5),
\end{aligned} \tag{4.9}$$

which is a convergent power series for  $|w+1| < \min\{1, |s+1|\}$ . Therefore, given contours  $C_{\pm}$  as above, if  $|w+1|$  is sufficiently small we may integrate term-by-term to obtain

$$\begin{aligned}
\frac{(-w)^{1/2}}{2\pi} \int_{C_+ \cup C_-} \frac{\theta_0(s) ds}{(-s)^{1/2}(s-w)} &= I_1 + \left[ I_2 - \frac{1}{2} I_1 \right] (w+1) + \left[ I_3 - \frac{1}{2} I_2 - \frac{1}{8} I_1 \right] (w+1)^2 \\
&+ \left[ I_4 - \frac{1}{2} I_3 - \frac{1}{8} I_2 - \frac{1}{16} I_1 \right] (w+1)^3 \\
&+ \left[ I_5 - \frac{1}{2} I_4 - \frac{1}{8} I_3 - \frac{1}{16} I_2 - \frac{5}{128} I_1 \right] (w+1)^4 + \mathcal{O}((w+1)^5),
\end{aligned} \tag{4.10}$$

where

$$I_k := \frac{1}{2\pi} \int_{C_+ \cup C_-} \frac{\theta_0(s) ds}{(-s)^{1/2}(s+1)^k}, \quad k = 1, \dots, 5. \tag{4.11}$$

By making the substitution  $\sigma = s^{-1}$ , we may rewrite these integrals in the form

$$I_k = -\frac{1}{2\pi} \int_{C_+ \cup C_-} \frac{\theta_0(\sigma) d\sigma}{(-\sigma)^{1/2}(\sigma+1)(\sigma^{-1}+1)^{k-1}}. \tag{4.12}$$

Therefore, we see immediately that

$$I_1 = -I_1 = 0, \tag{4.13}$$

while by averaging (4.11) and (4.12) we obtain for  $I_2$ ,  $I_3$ , and  $I_4$  the following:

$$I_2 = I_3 = \frac{1}{4\pi} \int_{C_+ \cup C_-} \frac{\theta_0(s)}{(-s)^{1/2}(s+1)} \frac{1-s}{1+s} ds \tag{4.14}$$

and

$$I_4 = \frac{1}{4\pi} \int_{C_+ \cup C_-} \frac{\theta_0(s)}{(-s)^{1/2}(s+1)} \frac{1-s^3}{(s+1)^3} ds. \tag{4.15}$$

Also, averaging  $I_5$  as given by (4.11) and (4.12) and comparing with (4.14) and (4.15) we obtain the identity

$$I_5 = 2I_4 - I_2 = 2I_4 - I_3. \tag{4.16}$$

Now we calculate  $I_2 = I_3$  exactly. Recalling the definitions (1.10) of the functions  $E(\cdot)$  and  $D(\cdot)$ , as well as the definition (1.14) of  $\theta_0(s)$ , we see that

$$I_2 = I_3 = -\frac{i}{8\pi} \int_{C_+ \cup C_-} \frac{\Psi(E(s))}{D(s)^3} E(s) E'(s) ds. \tag{4.17}$$

Since  $D(s)$  can be eliminated in favor of  $E(s)$  by the identity  $D(s)^2 = E(s)^2 + \frac{1}{4}$ , we would like to introduce  $v = -4iE(s)$  as a new integration variable. The two contours  $C_{\pm}$  are mapped under  $-4iE(\cdot)$  to two oppositely-oriented copies of the same contour. In particular,  $-4iE(C_-)$  is a teardrop-shaped contour beginning and ending at the point  $v = -4iE(\mathfrak{a}) = -4iE(\mathfrak{b})$  lying on the real axis to the right of  $v = 2$  and encircling the point  $v = 2 = -4iE(-1)$  once in the positive sense. Along the contours  $C_{\pm}$  we have the

identities  $D(s) = \pm(E(s)^2 + \frac{1}{4})^{1/2}$  (principal branch). These considerations show that both contributions from  $C_+$  and from  $C_-$  are equal, and so

$$I_2 = I_3 = \frac{i}{2\pi} \int_{-4iE(C_-)} \frac{\Psi(iv/4)}{(4-v^2)^{3/2}} (-2v) dv. \quad (4.18)$$

Next, we integrate by parts, using the fact that  $\Psi(E(\mathbf{a})) = 0$ :

$$I_2 = I_3 = \frac{i}{\pi} \int_{-4iE(C_-)} \frac{\varphi(v) dv}{(4-v^2)^{1/2}}, \quad \varphi(v) := \frac{d}{dv} \Psi(iv/4). \quad (4.19)$$

Collapsing the contour to the top and bottom of the interval  $(2, -4iE(\mathbf{a}))$  and noting that  $-4iE(\mathbf{a}) = -G(0)$  we obtain

$$I_2 = I_3 = \frac{2}{\pi} \int_2^{-G(0)} \frac{\varphi(v) dv}{\sqrt{v^2 - 4}}. \quad (4.20)$$

Now substituting from the definition (1.8) and exchanging the order of integration (see [7, Proposition 1.1]) leads to the identity

$$I_2 = I_3 = -\frac{1}{2} G^{-1}(-2) = -\frac{1}{2} x_{\text{crit}}. \quad (4.21)$$

Next we will show that

$$I_4 = \nu - \frac{1}{2} x_{\text{crit}} \quad (4.22)$$

where  $\nu > 0$  is defined by (1.30). Using (4.21) together with (4.14)–(4.15) and the definitions (1.10) of  $E(\cdot)$  and  $D(\cdot)$  gives

$$I_4 + \frac{1}{2} x_{\text{crit}} = I_4 - I_2 = \frac{i}{128\pi} \int_{C_+ \cup C_-} \frac{\Psi(E(s))}{D(s)^5} E(s) E'(s) ds. \quad (4.23)$$

Introducing  $v = -4iE(s)$  as a new integration variable, as above, yields

$$I_4 + \frac{1}{2} x_{\text{crit}} = \frac{1}{2\pi i} \int_{-4iE(C_-)} \frac{\Psi(iv/4)}{(4-v^2)^{5/2}} (-2v) dv, \quad (4.24)$$

and then integrating by parts,

$$I_4 + \frac{1}{2} x_{\text{crit}} = \frac{1}{3\pi i} \int_{-4iE(C_-)} \frac{\varphi(v) dv}{(4-v^2)^{3/2}}. \quad (4.25)$$

Note that by the substitution  $m = G(s)^2$ ,  $\varphi(v)$  may be written in the form

$$\varphi(v) = -\frac{v}{2} \int_{v^2}^{G(0)^2} \frac{\mathcal{G}(m)}{\sqrt{m-v^2} \sqrt{G(0)^2-m}} \frac{dm}{m}, \quad 0 < v < -G(0), \quad (4.26)$$

where the function  $\mathcal{G}$  is defined in terms of the initial condition  $G(\cdot)$  by (1.6). We need to write this formula in a way that admits analytic continuation to complex  $v$  such as those  $v \in -4iE(C_-)$ . To this end, let

$$T(m, z) := i(G(0)^2 - m)^{-1} \left( 1 - \frac{z - G(0)^2}{m - G(0)^2} \right)^{-1/2} \quad (4.27)$$

where the principal branch is meant, and then note that

$$\varphi(v) = \frac{v}{4} \oint_O T(m, v^2) \frac{\mathcal{G}(m)}{m} dm, \quad (4.28)$$

where  $O$  is a loop contour surrounding, in the negative (clockwise) sense, the straight-line branch cut of  $T(m, v^2)$  (viewed as a function of  $m$ ) connecting  $v^2$  with  $G(0)^2$ . In writing the formula (4.28) we therefore are using Assumption 1.4 to guarantee analyticity of  $\mathcal{G}$  in a neighborhood of  $m = G(0)^2$ . With the loop contour  $O$  fixed, it is clear that  $\varphi(v)$  as given by (4.28) is analytic if  $z = v^2$  varies in the region enclosed by  $O$ . In particular, we will assume that  $C_-$  has been chosen so that  $-4iE(C_-)$  is completely contained in this region. Since  $-4iE(C_-)$  and  $L$  are both fixed contours, we may exchange the order of integration upon substituting (4.28) into (4.25):

$$I_4 + \frac{1}{2} x_{\text{crit}} = \frac{1}{24\pi i} \oint_O \frac{\mathcal{G}(m)}{m} j(m) dm, \quad (4.29)$$



where the inner integral is now:

$$j(m) := \int_{-4iE(C_-)} \frac{T(m, v^2)}{(4 - v^2)^{3/2}} (2v) dv = \int_{[-4iE(C_-)]^2} \frac{T(m, z) dz}{(4 - z)^{3/2}}. \quad (4.30)$$

Note that the contour  $[-4iE(C_-)]^2$  is a teardrop-shaped contour beginning and ending at  $z = G(0)^2$  and encircling the part of the (principal) branch cut of  $(4 - z)^{3/2}$  between  $z = 4$  and  $z = G(0)^2 > 4$  once in the positive sense, and each  $m \in O$  lies outside of this closed contour. To evaluate  $j(m)$  we note that the branch cut of  $T(m, z)$  viewed as a function of  $z$  is the ray from  $z = m$  to  $z = \infty$  in the direction away from  $z = G(0)^2$ . By simple contour deformations of  $[-4iE(C_-)]^2$  taking into account that the integrand  $T(m, z)(4 - z)^{-3/2}$  is integrable at  $z = \infty$  and changes sign across all branch cuts, we obtain, for  $\Im\{m\} \neq 0$ ,

$$j(m) = 2 \operatorname{sgn}(\Im\{m\}) \int_{G(0)^2}^m \frac{T(m, z) dz}{(4 - z)^{3/2}}, \quad (4.31)$$

where the path of integration is a straight line in the region of analyticity of the integrand. Since  $\operatorname{sgn}(\Im\{m\})$  and  $\operatorname{sgn}(\Im\{z\})$  coincide, we may absorb the sign by changing the branch of  $(4 - z)^{3/2}$  in the lower-half  $z$ -plane:

$$j(m) = 2 \int_{G(0)^2}^m \frac{T(m, z) dz}{i(z - 4)^{3/2}}. \quad (4.32)$$

With the substitution

$$\alpha := \left(1 - \frac{z - G(0)^2}{m - G(0)^2}\right)^{1/2}, \quad 0 < \alpha < 1, \quad (4.33)$$

the formula for  $j(m)$  becomes

$$\begin{aligned} j(m) &= -4 \int_0^1 (m - 4 - (m - G(0)^2)\alpha^2)^{-3/2} d\alpha \\ &= \frac{4}{4 - m} \int_0^1 \frac{d}{d\alpha} \left[ \alpha (m - 4 - (m - G(0)^2)\alpha^2)^{-1/2} \right] d\alpha \\ &= \frac{4}{(4 - m)\sqrt{G(0)^2 - 4}}. \end{aligned} \quad (4.34)$$

Inserting this formula into (4.29) and taking into account that the only singularity of the integrand enclosed by the contour  $O$  is the simple pole at  $m = 4$  yields

$$I_4 + \frac{1}{2}x_{\text{crit}} = \frac{\mathcal{G}(4)}{12\sqrt{G(0)^2 - 4}}. \quad (4.35)$$

Evaluating  $\mathcal{G}(4)$  directly using the definition (1.6) then proves that indeed  $I_4 + \frac{1}{2}x_{\text{crit}} = \nu$  where  $\nu > 0$  is given by (1.30).

Using (4.13), (4.16), (4.21), and (4.22) in (4.10), and combining the resulting expansion with the expansions (4.8) and the analytic formula (4.7) for  $k(w; x, t)$  completes the proof of the Proposition.  $\square$

**4.3. Conformal mapping near  $w = -1$  and new spacetime coordinates.** From (4.4) it is clear that exactly at criticality,  $k(w) = \nu(w + 1)^3 + \mathcal{O}((w + 1)^4)$ . For small  $|\Delta x|$  and  $|t|$ , the cubic degeneration unfolds, with one real and two complex conjugate roots, or three real roots, near  $w = -1$ . The double critical point (double root of  $k'(w)$ ) unfolds generically to a pair  $w = w_{\pm}(x, t)$  of simple critical points near  $w = -1$ , and either these are both real or they form a complex-conjugate pair (in both cases since  $k(w) = k(w^*)^*$ ,  $(k(w_+) - k(w_-))^2 \in \mathbb{R}$ ).

The main idea here is that to unfold the cubic degeneracy it is really only necessary to take into account the lower-order terms in the Taylor expansion, and therefore it seems desirable to somehow replace  $k(w; x, t)$  by an appropriate cubic polynomial with coefficients depending on  $(x, t)$ . This issue has arisen frequently in the construction of local parametrices for matrix Riemann-Hilbert problems corresponding to certain *double-scaling limits*. Perhaps the first time such a replacement was made rigorous was in the paper of Baik, Deift, and Johansson [2], in which a certain double-scaling limit of orthogonal polynomials on the unit circle is analyzed, and the authors construct a local parametrix by essentially truncating an analogue of the Taylor expansion (4.4) after the cubic term. Of course such a truncation is not exact, so there are errors

incurred in modeling the jump matrices by others having cubic exponents, and the effect of these errors must be carefully controlled. Also, any artificial truncation of a Taylor series can only be accurate if the local parametrix is constructed in a disk centered at the expansion point whose radius tends to zero at some rate tied to the large parameter in the problem. This fact further implies that estimates must be supplied for norms of singular integral operators that are independent of the moving contour and it also typically means that divergence of the outer parametrix near the expansion point can lead to a sub-optimal mismatch with the local parametrix, leading in turn to sub-optimal estimates of errors. A significant advance was made by Claeys and Kuijlaars in [12] (see in particular section 5.6 of that paper); here the authors consider a similar double-scaling limit and take the approach of constructing a certain nontrivial conformal mapping  $W = W(w)$  of a fixed neighborhood of the expansion point to a neighborhood of the origin. The conformal mapping is more-or-less explicit, and it does not depend on the parameters (analogues of  $x$  and  $t$ ) driving the system to criticality. However, the point is that in a full fixed neighborhood of the expansion point, the analogue of the function  $k(w; x, t)$  is represented exactly (no truncation required) as a cubic polynomial in  $W$ . The difficulty that remains with this approach is that the coefficients of the cubic in  $W$  actually depend on  $w$  (so in fact it is not really a cubic after all). The remarkable approach of [12] is to simply substitute the  $w$ -dependent coefficients into a known solution formula for exactly cubic exponents (involving solutions of the linear differential equation whose isomonodromy deformations are governed by solutions of the Painlevé-II equation), resulting in a local parametrix in terms of Painlevé transcendents depending (through the coefficients of the cubic) on  $w$ . Confirming that such an approach actually provides a usable local parametrix requires exploiting *a priori* knowledge of the behavior of solutions of the nonlinear Painlevé-II equation and its auxiliary linear differential equation. If this information is available in a convenient form, the approach of Claeys and Kuijlaars delivers a vast improvement over earlier methods because it really works in a neighborhood of fixed size centered at the expansion point where critical points coalesce at criticality. The technique advanced in [12] has more recently been applied to problems of nonlinear wave theory as well (see, for example, [10]).

We choose instead to replace  $k(w; x, t)$  by a cubic in an exact way, an approach that provides all of the accuracy of the Claeys-Kuijlaars method but seems simpler and requires no *a priori* knowledge of the behavior of solutions of the local parametrix Riemann-Hilbert problem. The approach we are going to explain now has also been used recently to study a different kind of double-scaling limit for a matrix Riemann-Hilbert problem in [5]. As part of a careful study of the asymptotic behavior of exponential integrals with exponent functions having coalescing critical points (to generalize the steepest descent or saddle point method), Chester, Friedman, and Ursell [9] showed how to construct a substitution that rendered the exponent function in the integrand in the exact form of a cubic polynomial. Their method applies in the current context to establish that, *because the coefficient  $\nu$  in the expansion (4.4) of  $k(w; x, t)$  is strictly positive*, there is a suitable choice of new spacetime coordinates  $r = r(x, t)$  and  $s = s(x, t)$  for which the relation

$$k(w; x, t) = W^3 + rW - s \quad (4.36)$$

defines an invertible conformal mapping  $W = W(w) = W(w; x, t)$  between  $w \in U$  and  $W \in W(U)$  that preserves the real axis, for  $(x, t)$  near enough to criticality. Moreover, the new coordinates  $r$  and  $s$  depend continuously on  $(x, t)$  near criticality. Unlike in the Claeys-Kuijlaars method [12], neither  $r$  nor  $s$  depends on  $w$  with the cost that the conformal mapping will now depend on  $(x, t)$ .

The new coordinates  $r = r(x, t)$  and  $s = s(x, t)$  are to be determined so that under (4.36) the critical points of the cubic on the right-hand side, namely  $W = \pm(-r/3)^{1/2}$ , correspond to the two critical points of  $k(w)$  near  $w = -1$  when  $\Delta x$  and  $t$  are sufficiently small. Evaluating (4.36) for  $w = w_{\pm}(x, t)$  and  $W = \pm(-r(x, t)/3)^{1/2}$ , one obtains the formulae

$$\begin{aligned} r(x, t) &= -3 \left( \frac{1}{16} [k(w_+(x, t); x, t) - k(w_-(x, t); x, t)]^2 \right)^{1/3} \\ s(x, t) &= -\frac{1}{2} [k(w_+(x, t); x, t) + k(w_-(x, t); x, t)], \end{aligned} \quad (4.37)$$

where in the formula for  $r$  the real cube root is meant, and hence both  $r$  and  $s$  are real. Moreover, it is possible to show that  $r(x, t)$  and  $s(x, t)$  are analytic functions of  $x$  and  $t$  near criticality, and have two-variable

Taylor expansions of the form:

$$\begin{aligned} r(x, t) &= \frac{1}{2\nu^{1/3}} \Delta x + \mathcal{O}(\Delta x^2, t\Delta x, t^2) \\ s(x, t) &= t + \mathcal{O}(\Delta x^2, t\Delta x, t^2). \end{aligned} \quad (4.38)$$

To see this, first we find a unique root  $w = -1 + a$  of  $k''(w; x, t)$  such that  $a = 0$  at criticality and such that  $a = a(x, t)$  is an analytic function of  $(x, t)$  near criticality. Indeed, from the Taylor expansion (4.4) and the fact that  $\nu \neq 0$  we see that the analytic implicit function theorem applies and we obtain

$$a(x, t) = -\frac{1}{12\nu} \Delta x + \frac{1}{24\nu} t + \mathcal{O}(\Delta x^2, t\Delta x, t^2). \quad (4.39)$$

Now write  $w = -1 + a(x, t) + z$ , and re-expand  $k(w; x, t)$  about  $z = 0$ . From (4.4) and the definition of  $a(x, t)$  we obtain the convergent power series expansion (note that there is no quadratic term)

$$k(w; x, t) = k_0(x, t) + k_1(x, t)z + \sum_{n=3}^{\infty} k_n(x, t)z^n, \quad (4.40)$$

where the coefficients are all analytic functions of  $(x, t)$  near criticality, and in particular,

$$k_0(x, t) = -t + \mathcal{O}(\Delta x^2, t\Delta x, t^2), \quad (4.41)$$

$$k_1(x, t) = \frac{1}{2} \Delta x + \mathcal{O}(\Delta x^2, t\Delta x, t^2), \quad (4.42)$$

$$k_3(x, t) = \nu + \mathcal{O}(\Delta x, t). \quad (4.43)$$

The radius of convergence of this series is bounded away from zero near criticality. The equation satisfied by the critical points of  $k$  is

$$k'(w; x, t) = k_1(x, t) + 3k_3(x, t)z^2 + \sum_{n=4}^{\infty} nk_n(x, t)z^{n-1} = 0. \quad (4.44)$$

Let  $\rho$  be any number satisfying the equation

$$\rho^2 = -\frac{k_1(x, t)}{3k_3(x, t)}. \quad (4.45)$$

Note that  $\rho^2$  is an analytic function of  $(x, t)$  near criticality because  $k_3(x_{\text{crit}}, 0) = \nu > 0$ . If  $\rho = 0$  then  $k_1(x, t) = 0$  and  $z = 0$  is a double root of (4.44). So we suppose that  $\rho \neq 0$ , and we rescale  $z$  by writing  $z = \rho q$  for some new unknown  $q$ . Dividing through by  $3k_3(x, t) \neq 0$  and canceling a factor of  $\rho^2$  then converts (4.44) into the equation

$$q^2 + \sum_{n=4}^{\infty} \frac{nk_n(x, t)}{3k_3(x, t)} \rho^{n-3} q^{n-1} = 1. \quad (4.46)$$

The coefficients of  $q^{n-1}$  in the sum are analytic functions of the three variables  $x, t$ , and  $\rho$ . When  $\rho = 0$  there are two solutions,  $q = \pm 1$ , and by the implicit function theorem there are two corresponding solutions for  $\rho \approx 0$  and  $(x, t)$  near criticality. The two solutions are related by the symmetry  $(\rho, q) \mapsto (-\rho, -q)$ . We may write them in the form

$$q = q_{\pm}(x, t, \rho) = \pm \left( 1 + \sum_{n=1}^{\infty} q_n(x, t) (\pm \rho)^n \right). \quad (4.47)$$

The coefficients  $q_n(x, t)$  are all analytic functions of  $(x, t)$  near criticality, and the radius of convergence of the series with respect to  $\rho$  is bounded away from zero near criticality. The corresponding critical points are written in terms of  $z$  as

$$z = z_{\pm}(x, t, \rho) = \pm \rho + \sum_{n=1}^{\infty} q_n(x, t) (\pm \rho)^{n+1}. \quad (4.48)$$

To calculate  $r(x, t)$  and  $s(x, t)$  we need to evaluate  $k(w; x, t)$  at the critical points. This is accomplished by substitution of the series (4.48) into (4.40):

$$\begin{aligned} k(-1 + a(x, t) + z_{\pm}(x, t, \rho); x, t) &= k_0(x, t) + k_1(x, t) \left( \pm \rho + \sum_{n=1}^{\infty} q_n(x, t) (\pm \rho)^{n+1} \right) \\ &\quad + \sum_{m=3}^{\infty} k_m(x, t) \left( \pm \rho + \sum_{n=1}^{\infty} q_n(x, t) (\pm \rho)^{n+1} \right)^m \\ &= k_0(x, t) \pm \rho^3 \sum_{n=0}^{\infty} g_n(x, t) (\pm \rho)^n, \end{aligned} \quad (4.49)$$

where we have used (4.45), and  $g_0(x, t) := -2k_3(x, t)$ . Here again, the coefficients  $g_n(x, t)$  are certain analytic functions of  $(x, t)$  near criticality, and the radius of convergence of the power series in  $\pm \rho$  is bounded below near criticality. Then, by definition, we have

$$\begin{aligned} r(x, t) &= -3 \left( \frac{1}{16} \left[ 2g_0(x, t) \rho^3 \left( 1 + \sum_{j=1}^{\infty} \frac{g_{2j}(x, t)}{g_0(x, t)} (\pm \rho)^{2j} \right) \right]^2 \right)^{1/3} \\ &= \frac{k_1(x, t)}{k_3(x, t)^{1/3}} \left( 1 - \frac{1}{2} \sum_{j=1}^{\infty} \frac{g_{2j}(x, t)}{k_3(x, t)} \left[ -\frac{k_1(x, t)}{3k_3(x, t)} \right]^j \right)^{2/3}, \end{aligned} \quad (4.50)$$

where we have used the definition of  $g_0(x, t)$  in terms of  $k_3(x, t)$  and (4.45). Therefore,  $r(x, t)$  is clearly an analytic function of  $(x, t)$  near criticality. Similarly, by definition we have

$$\begin{aligned} s(x, t) &= -\frac{1}{2} \left[ 2k_0(x, t) + 2 \sum_{j=0}^{\infty} g_{2j+1}(x, t) (\pm \rho)^{2j+4} \right] \\ &= -k_0(x, t) - \sum_{j=0}^{\infty} g_{2j+1}(x, t) \left( -\frac{k_1(x, t)}{3k_3(x, t)} \right)^{j+2}, \end{aligned} \quad (4.51)$$

which again is obviously an analytic function of  $(x, t)$  near criticality. The leading terms of  $r(x, t)$  and  $s(x, t)$  near criticality are easy to calculate from these formulae, with the result being (4.38).

As  $W(\cdot; x, t)$  is a conformal map, its inverse is an analytic function mapping a neighborhood of  $W = 0$  to a neighborhood of  $w = -1$ . We now define

$$w_* = w_*(x, t) := W^{-1}(0; x, t), \quad (4.52)$$

which is a real analytic function of  $x$  and  $t$  near criticality. Note also that  $W'(w_*(x, t); x, t) > 0$  (prime denotes differentiation with respect to  $w$ ) near criticality. It is not difficult to obtain the Taylor expansion of the conformal mapping  $W(w; x, t)$  about  $w = w_*$  exactly at criticality. Indeed, from (4.4) and (4.36) with  $\Delta x = t = r = s = 0$ , we obtain the equation

$$W(w)^3 = \nu(w+1)^3 \left[ 1 + \frac{3}{2}(w+1) + \mathcal{O}((w+1)^2) \right], \quad \text{at criticality}, \quad (4.53)$$

and analytically blowing up the cubic degeneracy we obtain

$$W(w) = \nu^{1/3}(w+1) \left[ 1 + \frac{1}{2}(w+1) + \mathcal{O}((w+1)^2) \right], \quad \text{at criticality}. \quad (4.54)$$

In particular, this implies that

$$W'(w_*) = W''(w_*) = \nu^{1/3} > 0, \quad \text{at criticality}. \quad (4.55)$$

Also, it is clear that  $W(-1; x_{\text{crit}}, 0) = 0$ , and therefore  $w_*(x_{\text{crit}}, 0) = -1$ . Since  $w_*(x, t)$  is an analytic function of  $x$  and  $t$ , it then follows that

$$w_*(x, t) = -1 + \mathcal{O}(\Delta x, t), \quad (4.56)$$

and furthermore,

$$W'(w_*(x, t); x, t) = \nu^{1/3} + \mathcal{O}(\Delta x, t) \quad \text{and} \quad W''(w_*(x, t); x, t) = \nu^{1/3} + \mathcal{O}(\Delta x, t) \quad (4.57)$$

near criticality.

Now let  $\zeta$  and  $y$  be scaled versions of  $W$  and  $r$  respectively:

$$\zeta := \frac{W}{\epsilon_N^{1/3}} \quad \text{and} \quad y := \frac{r}{\epsilon_N^{2/3}}. \quad (4.58)$$

We will regard  $y$  as being bounded. It then follows that *without approximation* the exponent appearing in the jump matrix  $\mathbf{V}_{\mathbf{P}}(w)$  takes the form of a simple cubic with a formally large constant term:

$$\frac{k(w; x, t)}{\epsilon_N} = \zeta^3 + y\zeta - \frac{s}{\epsilon_N}. \quad (4.59)$$

**4.4. Formulation of an inner model problem.** To begin with, we wish to find a simpler representation for  $\dot{\mathbf{P}}_m^{\text{out}}(w)$  valid when  $w$  is close to  $w = w_* \approx -1$ . Using (4.58), we write  $\dot{\mathbf{P}}_m^{\text{out}}(w)$  in the form

$$\dot{\mathbf{P}}_m^{\text{out}}(w) = \eta_m(w)^{\sigma_3} (-W(w))^{(m-1/2)\sigma_3} = \eta_m(w)^{\sigma_3} \epsilon_N^{(2m-1)\sigma_3/6} (-\zeta)^{(2m-1)\sigma_3/2}, \quad (4.60)$$

where  $\eta_m(w)$  is independent of  $\epsilon_N$  and is analytic and nonvanishing in a neighborhood  $U$  of  $w = w_*$ :

$$\eta_m(w) := (\sqrt{-w} + \sqrt{-w_*})^{1-2m} \left( \frac{W(w)}{w - w_*} \right)^{1/2-m}. \quad (4.61)$$

(Analyticity follows since from (4.52) we have  $W(w_*) = 0$  and  $W$  is analytic at  $w_*$  with  $W'(w_*) > 0$ .)

Given a value of  $m \in \mathbb{Z}$  (to be determined below), we wish to construct an inner model, valid for  $w \in U$ , of the matrix  $\tilde{\mathbf{P}}(w)$  (related to  $\mathbf{P}(w)$  for  $w \in U$  via (4.2)). We temporarily denote this model as  $\mathbf{W}(\zeta(w))$ , and we want it to have the following properties:

- Supposing that in  $U$  the lens boundaries are identified with arcs of the curves  $\arg(W(w)) = \pm\pi/3$  and  $\arg(W(w)) = \pm 2\pi/3$ , which makes them segments (of length proportional to  $\epsilon_N^{-1/3}$ ) of straight rays in the  $\zeta$ -plane, the inner model  $\mathbf{W}(\zeta(w))$  should be analytic exactly where  $\tilde{\mathbf{P}}(w)$  is within  $U$  and should satisfy *exactly the same jump conditions* that  $\tilde{\mathbf{P}}(w)$  does within  $U$ .
- The inner model  $\mathbf{W}(\zeta)$  should match onto the latter factors in (4.60) along the disc boundary  $\partial U$  in the sense that  $\mathbf{W}(\zeta)$  may be analytically continued from each sector of the domain  $\zeta \in \epsilon_N^{-1/3} W(U)$  to the corresponding infinite sector in the  $\zeta$ -plane, and that

$$\lim_{\zeta \rightarrow \infty} \mathbf{W}(\zeta) \epsilon_N^{(1-2m)\sigma_3/6} (-\zeta)^{(1-2m)\sigma_3/2} = \mathbb{I}, \quad (4.62)$$

with the limit being uniform with respect to direction in each of the six sectors of analyticity. Since  $\partial U$  is fixed and bounded away from  $w_* \approx -1$ , upon scaling its image under  $W(\cdot; x, t)$  by  $\epsilon_N^{-1/3}$  to work in terms of the variable  $\zeta$ , we see that  $w \in \partial U$  corresponds to  $\zeta \rightarrow \infty$  at a uniform rate of  $\epsilon_N^{-1/3}$ .

Since  $\mathbf{W}(\zeta(w))$  depends on  $w$  through  $\zeta(w)$ , it is convenient to use (4.59) to write the jump matrices for  $\tilde{\mathbf{P}}(w)$  (and hence also for  $\mathbf{W}(\zeta(w))$ ) in terms of  $\zeta$ , which shows that the jump matrices involve the product of exponentials  $e^{\mp s/\epsilon_N} e^{\pm(\zeta^3 + y\zeta)}$ . The constant ( $\zeta$ -independent) factors  $e^{\mp s/\epsilon_N}$  present in the jump matrices can be removed, and simultaneously the presence of  $\epsilon_N$  in the normalization condition (4.62) can be eliminated, by defining the equivalent unknown

$$\mathbf{Z}_m(\zeta; y) := \epsilon_N^{(1-2m)\sigma_3/6} e^{-\frac{1}{2}s\sigma_3/\epsilon_N} \mathbf{W}(\zeta) e^{\frac{1}{2}s\sigma_3/\epsilon_N}. \quad (4.63)$$

The conditions previously discussed as being desirable for  $\mathbf{W}(\zeta)$  are then easily seen to be equivalent to those of the following problem for  $\mathbf{Z}_m(\zeta; y)$ :

**Riemann-Hilbert Problem 4.1** (Inner model problem near criticality). *Let a real number  $y$  and an integer  $m$  be fixed. Seek a matrix  $\mathbf{Z}_m(\zeta; y)$  with the following properties:*

**Analyticity:**  $\mathbf{Z}_m(\zeta; y)$  is analytic in  $\zeta$  except along the rays  $\arg(\zeta) = k\pi/3$ ,  $k = 0, \dots, 5$ , from each sector of analyticity it may be continued to a slightly larger sector, and in each sector is Hölder continuous up to the boundary in a neighborhood of  $\zeta = 0$ .

**Jump condition:** The jump conditions satisfied by the matrix function  $\mathbf{Z}_m(\zeta; y)$  are of the form  $\mathbf{Z}_{m+}(\zeta; y) = \mathbf{Z}_{m-}(\zeta; y) \mathbf{V}_\mathbf{Z}(\zeta; y)$ , where the jump matrix  $\mathbf{V}_\mathbf{Z}(\zeta; y)$  is as shown in Figure 4.2 and all

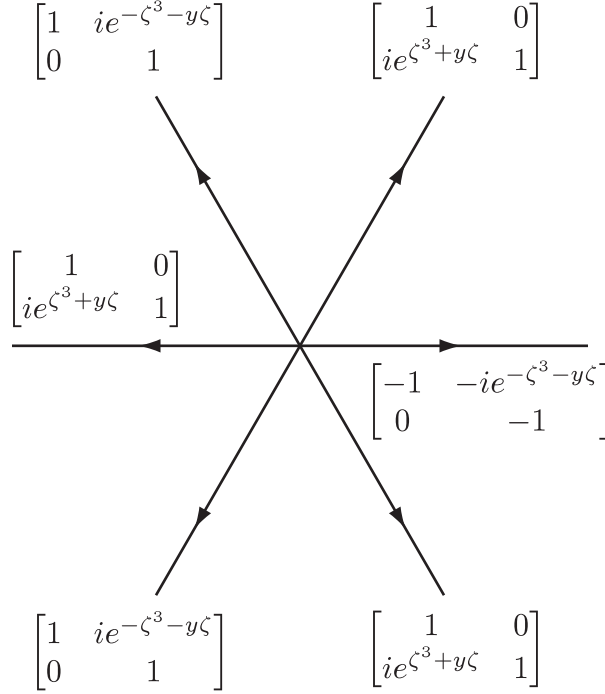


FIGURE 4.2. The jump matrix  $\mathbf{V}_\mathbf{Z}(\zeta; y)$ .

rays are oriented toward infinity.

**Normalization:** The matrix  $\mathbf{Z}_m(\zeta; y)$  satisfies the condition

$$\lim_{\zeta \rightarrow \infty} \mathbf{Z}_m(\zeta; y) (-\zeta)^{(1-2m)\sigma_3/2} = \mathbb{I}, \quad (4.64)$$

with the limit being uniform with respect to direction in each of the six sectors of analyticity.

Suppose that Riemann-Hilbert Problem 4.1 has a unique solution for  $y \in (y_-, y_+) \subset \mathbb{R}$  and for some integer  $m$ . We will now describe how to use it to create a model for  $\mathbf{P}(w)$  valid in the neighborhood  $U$  assuming that  $y \in (y_-, y_+)$ ; we set

$$\begin{aligned} \dot{\mathbf{P}}_m^{\text{in}}(w) &:= \eta_m(w)^{\sigma_3} \mathbf{W}(\zeta(w)) d_N(w)^{-\sigma_3} \\ &= \eta_m(w)^{\sigma_3} \epsilon_N^{(2m-1)\sigma_3/6} e^{\frac{1}{2}s\sigma_3/\epsilon_N} \mathbf{Z}_m(\zeta(w); y) e^{-\frac{1}{2}s\sigma_3/\epsilon_N} d_N(w)^{-\sigma_3}, \quad w \in U. \end{aligned} \quad (4.65)$$

The effect of post-multiplication by  $d_N(w)^{-\sigma_3}$  is simply to restore the factors involving  $T_N(w)$  to the jump conditions. We emphasize that within the fixed neighborhood  $U$  of  $w = -1$ , the inner model matrix  $\dot{\mathbf{P}}_m^{\text{in}}(w)$  satisfies exactly the same jump conditions along the six contour arcs meeting at  $w = w_*$  as does  $\mathbf{P}(w)$ . It appears reasonable to propose a global model for  $\mathbf{P}(w)$  in the following form:

$$\dot{\mathbf{P}}_m(w) := \begin{cases} \dot{\mathbf{P}}_m^{\text{in}}(w), & w \in U \\ \dot{\mathbf{P}}_m^{\text{out}}(w), & w \notin \overline{U}. \end{cases} \quad (4.66)$$

The integer  $m$  is evidently at our disposal. We will later describe how it should be chosen.

## 5. SOLUTION OF THE INNER MODEL PROBLEM

**5.1. Symmetry analysis of the inner model problem.** While it may not yet be clear whether there exists a solution of Riemann-Hilbert Problem 4.1 for any  $y \in \mathbb{R}$  and  $m \in \mathbb{Z}$  at all, it is a standard argument that there exists at most one solution for each  $y \in \mathbb{R}$  and  $m \in \mathbb{Z}$ , and that every solution must satisfy  $\det(\mathbf{Z}_m(\zeta; y)) \equiv 1$ . In this short section we suppose that  $y \in \mathbb{R}$  and  $m \in \mathbb{Z}$  are values for which Riemann-Hilbert Problem 4.1 has a (unique and unimodular) solution, and we examine some of the consequences of the existence. We will later show that  $\mathbf{Z}_m(\zeta; y)(-\zeta)^{(1-2m)\sigma_3/2}$  has a complete asymptotic expansion in descending powers of  $\zeta$  as  $\zeta \rightarrow \infty$ :

$$\mathbf{Z}_m(\zeta; y)(-\zeta)^{(1-2m)\sigma_3/2} = \mathbb{I} + \mathbf{A}_m(y)\zeta^{-1} + \mathbf{B}_m(y)\zeta^{-2} + \mathbf{C}_m(y)\zeta^{-3} + \mathcal{O}(\zeta^{-4}), \quad \zeta \rightarrow \infty \quad (5.1)$$

and where the matrix coefficients are the same regardless of which of the six sectors of analyticity of  $\mathbf{Z}_m(\zeta; y)$  is used to compute the expansion.

Given a solution  $\mathbf{Z}_m(\zeta; y)$  of Riemann-Hilbert Problem 4.1, consider the related matrix

$$\mathbf{F}(\zeta) := \mathbf{Z}_m(\zeta^*; y)^*. \quad (5.2)$$

Since the jump contour for  $\mathbf{Z}_m(\zeta; y)$  is invariant under complex conjugation,  $\mathbf{F}(\zeta)$  will be analytic in the same domain that  $\mathbf{Z}_m(\zeta; y)$  is. It is a direct calculation that  $\mathbf{F}(\zeta)$  satisfies exactly the same jump conditions as does  $\mathbf{Z}_m(\zeta; y)$ . Therefore, the matrix  $\mathbf{F}(\zeta)\mathbf{Z}_m(\zeta; y)^{-1}$  extends continuously to the jump contour from both sides of each arc. It follows from the classical sense in which the boundary values are attained (even at the self-intersection point  $\zeta = 0$ ), and from the fact that  $\mathbf{Z}_m(\zeta; y)$  is unimodular, that in fact  $\mathbf{F}(\zeta)\mathbf{Z}_m(\zeta; y)^{-1}$  is an entire function of  $\zeta$ . Moreover,

$$\begin{aligned} \lim_{\zeta \rightarrow \infty} \mathbf{F}(\zeta)\mathbf{Z}_m(\zeta; y)^{-1} &= \lim_{\zeta \rightarrow \infty} \mathbf{F}(\zeta)(-\zeta)^{(1-2m)\sigma_3/2} \cdot \left[ \mathbf{Z}_m(\zeta; y)(-\zeta)^{(1-2m)\sigma_3/2} \right]^{-1} \\ &= \lim_{\zeta \rightarrow \infty} \mathbf{Z}_m(\zeta^*; y)^*(-\zeta)^{(1-2m)\sigma_3/2} = \lim_{\zeta \rightarrow \infty} \left[ \mathbf{Z}_m(\zeta^*; y)(-\zeta^*)^{(1-2m)\sigma_3/2} \right]^* = \mathbb{I}, \end{aligned} \quad (5.3)$$

so by Liouville's Theorem,  $\mathbf{Z}_m(\zeta^*; y)^* = \mathbf{F}(\zeta) = \mathbf{Z}_m(\zeta; y)$ . Applying this symmetry to the expansion (5.1) shows that the elements of the matrices  $\mathbf{A}_m(y)$  and  $\mathbf{B}_m(y)$  (and in fact all of the matrix expansion coefficients) are real numbers.

We can obtain more detailed information by exploiting a further symmetry in the special case of  $m = 0$ . Indeed, let us compare  $\mathbf{Z}_0(\zeta; y)$  with the matrix

$$\mathbf{F}(\zeta) := \begin{cases} \mathbf{Z}_0(-\zeta; y)i\sigma_1, & \Im\{\zeta\} > 0 \\ -\mathbf{Z}_0(-\zeta; y)i\sigma_1, & \Im\{\zeta\} < 0. \end{cases} \quad (5.4)$$

Clearly,  $\det(\mathbf{F}(\zeta)) \equiv 1$ , and it is a direct matter to check that  $\mathbf{F}(\zeta)$  is analytic precisely where  $\mathbf{Z}_0(\zeta; y)$  is, and  $\mathbf{F}(\zeta)$  satisfies exactly the same jump conditions on the six rays of Riemann-Hilbert Problem 4.1 as does  $\mathbf{Z}_0(\zeta; y)$ . Therefore,  $\mathbf{F}(\zeta)\mathbf{Z}_0(\zeta; y)^{-1}$  is an entire unimodular matrix function of  $\zeta$ . To identify this entire function as a polynomial, it is enough to extract the non-decaying terms in its asymptotic expansion as  $\zeta \rightarrow \infty$ ; Assuming without loss of generality that  $\Im\{\zeta\} > 0$  so that  $(-\zeta)^{\sigma_3/2} = -i\sigma_3\zeta^{\sigma_3/2}$ , we have

$$\begin{aligned} \mathbf{F}(\zeta)\mathbf{Z}_0(\zeta; y)^{-1} &= \mathbf{F}(\zeta)(-i\sigma_3)\zeta^{\sigma_3/2} \cdot \left[ \mathbf{Z}_0(\zeta; y)(-\zeta)^{\sigma_3/2} \right]^{-1} \\ &= \mathbf{Z}_0(-\zeta; y)\sigma_1\sigma_3\zeta^{\sigma_3/2} \cdot (\mathbb{I} - \mathbf{A}_0(y)\zeta^{-1} + \mathcal{O}(\zeta^{-2})) \\ &= \mathbf{Z}_0(-\zeta; y)\zeta^{\sigma_3/2} \begin{bmatrix} 0 & -\zeta^{-1} \\ \zeta & 0 \end{bmatrix} (\mathbb{I} - \mathbf{A}_0(y)\zeta^{-1} + \mathcal{O}(\zeta^{-2})) \\ &= (\mathbb{I} - \mathbf{A}_0(y)\zeta^{-1} + \mathcal{O}(\zeta^{-2})) \begin{bmatrix} 0 & -\zeta^{-1} \\ \zeta & 0 \end{bmatrix} (\mathbb{I} - \mathbf{A}_0(y)\zeta^{-1} + \mathcal{O}(\zeta^{-2})) \\ &= \sigma_-\zeta - \mathbf{A}_0(y)\sigma_- - \sigma_-\mathbf{A}_0(y) + \mathcal{O}(\zeta^{-1}), \quad \zeta \rightarrow \infty, \end{aligned} \quad (5.5)$$

where

$$\sigma_+ := \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad \sigma_- := \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad (5.6)$$

and so  $\mathbf{F}(\zeta)\mathbf{Z}_0(\zeta; y)^{-1}$  may be identified with the polynomial terms in this expansion; therefore we arrive at the identity

$$\mathbf{F}(\zeta)\mathbf{Z}_0(\zeta; y)^{-1} = \begin{bmatrix} -A_{0,12}(y) & 0 \\ \zeta - A_{0,11}(y) - A_{0,22}(y) & -A_{0,12}(y) \end{bmatrix}. \quad (5.7)$$

Since both sides represent unimodular matrices, we learn that

$$A_{0,12}(y)^2 = 1. \quad (5.8)$$

We will later obtain a refinement of (5.8), namely that  $A_{0,12}(y) = 1$  holds for all real  $y$ . The sign information will follow from the explicit solution of Riemann-Hilbert Problem 4.1 for  $m = 0$  in terms of special functions, which will be possible precisely as a consequence of the identity (5.8).

**5.2. Differential equations derived from the inner model problem. Lax equations and Painlevé-II.** The only parameters in the Riemann-Hilbert problem characterizing  $\mathbf{Z}_m(\zeta; y)$  are  $m \in \mathbb{Z}$  and  $y \in \mathbb{R}$ . Assuming existence for some fixed  $m \in \mathbb{Z}$  and for  $y$  in some open set, we will now investigate some consequences of the dependence of the solution on  $y$ . We suppose that the expansion (5.1) holds in the stronger sense that the following are also true:

$$\frac{\partial}{\partial y} \left[ \mathbf{Z}_m(\zeta; y)(-\zeta)^{(1-2m)\sigma_3/2} \right] = \mathcal{O}(\zeta^{-1}) \quad \text{and} \quad \frac{\partial}{\partial \zeta} \left[ \mathbf{Z}_m(\zeta; y)(-\zeta)^{(1-2m)\sigma_3/2} \right] = \mathcal{O}(\zeta^{-2}), \quad \zeta \rightarrow \infty. \quad (5.9)$$

The matrix  $\mathbf{L}_m(y, \zeta)$  defined by

$$\mathbf{L}_m(y, \zeta) := \mathbf{Z}_m(\zeta; y)e^{-(\zeta^3 + y\zeta)\sigma_3/2} \quad (5.10)$$

has jump discontinuities along the six rays mediated by jump matrices that are independent of both  $y$  and  $\zeta$ . Given the compatibility of the jump matrices at  $\zeta = 0$ , this implies that the matrices

$$\mathbf{U}_m(y, \zeta) := \frac{\partial \mathbf{L}_m}{\partial y} \mathbf{L}_m^{-1} \quad \text{and} \quad \mathbf{V}_m(y, \zeta) := \frac{\partial \mathbf{L}_m}{\partial \zeta} \mathbf{L}_m^{-1} \quad (5.11)$$

are both entire functions of  $\zeta$ . Using the above expansions of  $\mathbf{Z}_m(\zeta; y)$  and its derivatives, we easily obtain the following expansions involving  $\mathbf{L}_m$ :

$$\begin{aligned} \frac{\partial \mathbf{L}_m}{\partial y} &= \left( -\frac{\partial \Theta}{\partial y} [\mathbb{I} + \mathbf{A}_m(y)\zeta^{-1} + \mathbf{B}_m(y)\zeta^{-2} + \mathcal{O}(\zeta^{-3})] \sigma_3 + \mathcal{O}(\zeta^{-1}) \right) e^{-\Theta\sigma_3} \\ &= \left( -\frac{1}{2}\sigma_3\zeta - \frac{1}{2}\mathbf{A}_m(y)\sigma_3 + \mathcal{O}(\zeta^{-1}) \right) e^{-\Theta\sigma_3} \\ \frac{\partial \mathbf{L}_m}{\partial \zeta} &= \left( -\frac{\partial \Theta}{\partial \zeta} [\mathbb{I} + \mathbf{A}_m(y)\zeta^{-1} + \mathbf{B}_m(y)\zeta^{-2} + \mathcal{O}(\zeta^{-3})] \sigma_3 + \mathcal{O}(\zeta^{-2}) \right) e^{-\Theta\sigma_3} \\ &= \left( -\frac{3}{2}\sigma_3\zeta^2 - \frac{3}{2}\mathbf{A}_m(y)\sigma_3\zeta - \frac{3}{2}\mathbf{B}_m(y)\sigma_3 - \frac{1}{2}y\sigma_3 + \mathcal{O}(\zeta^{-1}) \right) e^{-\Theta\sigma_3} \\ \mathbf{L}_m^{-1} &= e^{\Theta\sigma_3} (\mathbb{I} - \mathbf{A}_m(y)\zeta^{-1} + [\mathbf{A}_m(y)^2 - \mathbf{B}_m(y)]\zeta^{-2} + \mathcal{O}(\zeta^{-3})), \end{aligned} \quad (5.12)$$

where

$$\Theta := \frac{1}{2} [\zeta^3 + y\zeta + (1 - 2m)\log(-\zeta)]. \quad (5.13)$$

We can then easily see that both  $\mathbf{U}_m$  and  $\mathbf{V}_m$  grow algebraically in  $\zeta$  as  $\zeta \rightarrow \infty$  and hence are necessarily polynomials:

$$\begin{aligned} \mathbf{U}_m(y, \zeta) &= -\frac{1}{2}\sigma_3\zeta + \frac{1}{2}[\sigma_3, \mathbf{A}_m(y)] \\ \mathbf{V}_m(y, \zeta) &= -\frac{3}{2}\sigma_3\zeta^2 + \frac{3}{2}[\sigma_3, \mathbf{A}_m(y)]\zeta - \frac{3}{2}[\sigma_3, \mathbf{A}_m(y)]\mathbf{A}_m(y) + \frac{3}{2}[\sigma_3, \mathbf{B}_m(y)] - \frac{1}{2}y\sigma_3. \end{aligned} \quad (5.14)$$

We may rewrite these in the form

$$\begin{aligned} \mathbf{U}_m(y, \zeta) &= -\frac{1}{2}\sigma_3\zeta + \begin{bmatrix} 0 & \mathcal{U}_m(y) \\ -\mathcal{V}_m(y) & 0 \end{bmatrix} \\ \mathbf{V}_m(y, \zeta) &= -\frac{3}{2}\sigma_3\zeta^2 + 3 \begin{bmatrix} 0 & \mathcal{U}_m(y) \\ -\mathcal{V}_m(y) & 0 \end{bmatrix} \zeta + \frac{1}{2} \begin{bmatrix} -6\mathcal{U}_m(y)\mathcal{V}_m(y) - y & 2\mathcal{W}_m(y) \\ -2\mathcal{Z}_m(y) & 6\mathcal{U}_m(y)\mathcal{V}_m(y) + y \end{bmatrix}, \end{aligned} \quad (5.15)$$



where

$$\begin{aligned}
\mathcal{U}_m(y) &:= A_{m,12}(y) \\
\mathcal{V}_m(y) &:= A_{m,21}(y) \\
\mathcal{W}_m(y) &:= 3B_{m,12}(y) - 3A_{m,12}(y)A_{m,22}(y) \\
\mathcal{Z}_m(y) &:= 3B_{m,21}(y) - 3A_{m,21}(y)A_{m,11}(y).
\end{aligned} \tag{5.16}$$

The matrix  $\mathbf{L}_m$  is therefore a simultaneous (and fundamental, since  $\det(\mathbf{L}_m) \equiv 1$ ) solution of the *Lax equations*

$$\frac{\partial \mathbf{L}_m}{\partial y} = \mathbf{U}_m(y, \zeta) \mathbf{L}_m \quad \text{and} \quad \frac{\partial \mathbf{L}_m}{\partial \zeta} = \mathbf{V}_m(y, \zeta) \mathbf{L}_m, \tag{5.17}$$

and therefore the compatibility condition

$$\frac{\partial \mathbf{U}_m}{\partial \zeta} - \frac{\partial \mathbf{V}_m}{\partial y} + [\mathbf{U}_m, \mathbf{V}_m] = 0 \tag{5.18}$$

holds identically in  $\zeta$ . Separating the powers of  $\zeta$  leads to the following system of differential equations (the *Painlevé-II system*) governing the quantities defined by (5.16):

$$\begin{aligned}
\mathcal{U}'_m(y) &= -\frac{1}{3}\mathcal{W}_m(y) \\
\mathcal{V}'_m(y) &= \frac{1}{3}\mathcal{Z}_m(y) \\
\mathcal{W}'_m(y) &= 6\mathcal{U}_m(y)^2\mathcal{V}_m(y) + y\mathcal{U}_m(y) \\
\mathcal{Z}'_m(y) &= -6\mathcal{U}_m(y)\mathcal{V}_m(y)^2 - y\mathcal{V}_m(y).
\end{aligned} \tag{5.19}$$

Although satisfied by quantities evidently depending on  $m \in \mathbb{Z}$ , the Painlevé-II system does not involve  $m$  in any explicit way. Eliminating  $\mathcal{W}_m$  and  $\mathcal{Z}_m$  yields the coupled system of second-order Painlevé-II-type equations

$$\mathcal{U}''_m(y) + 2\mathcal{U}_m(y)^2\mathcal{V}_m(y) + \frac{1}{3}y\mathcal{U}_m(y) = 0 \quad \text{and} \quad \mathcal{V}''_m(y) + 2\mathcal{U}_m(y)\mathcal{V}_m(y)^2 + \frac{1}{3}y\mathcal{V}_m(y) = 0. \tag{5.20}$$

In fact, functions related in an elementary way to  $\mathcal{U}_m$  and  $\mathcal{V}_m$  turn out to satisfy uncoupled second-order equations of Painlevé-II type, but we must use the *inhomogeneous* form of the Painlevé-II equation, and the coefficient of inhomogeneity will depend on  $m \in \mathbb{Z}$ . To see that this is so, we follow [18, pages 154–155] by multiplying the first equation of (5.20) by  $\mathcal{V}_m$ , the second by  $\mathcal{U}_m$ , and subtracting:

$$0 = \mathcal{V}_m(y)\mathcal{U}''_m(y) - \mathcal{U}_m(y)\mathcal{V}''_m(y) = \frac{d}{dy} [\mathcal{V}_m(y)\mathcal{U}'_m(y) - \mathcal{U}_m(y)\mathcal{V}'_m(y)]. \tag{5.21}$$

Hence, the quantity

$$\lambda_m := \mathcal{V}_m(y)\mathcal{U}'_m(y) - \mathcal{U}_m(y)\mathcal{V}'_m(y) = -\frac{1}{3} [\mathcal{V}_m(y)\mathcal{W}_m(y) + \mathcal{U}_m(y)\mathcal{Z}_m(y)] \tag{5.22}$$

is a constant, independent of  $y$ .

To compute the value of  $\lambda_m$  for the particular solution of (5.19) corresponding to the specific Riemann-Hilbert problem for  $\mathbf{Z}_m(\zeta; y)$ , we first consider the direct problem for the Lax equation  $\mathbf{L}_{m,\zeta} = \mathbf{V}_m \mathbf{L}_m$ , formally expanding  $\mathbf{L}_m(y, \zeta)$  for large  $\zeta$  by assuming the form

$$\mathbf{L}_m(y, \zeta) \sim \left( \mathbb{I} + \sum_{k=1}^{\infty} \mathbf{Y}_{m,k}(y) \zeta^{-k} \right) e^{\mathbf{D}_m(y, \zeta)}, \quad \zeta \rightarrow \infty, \tag{5.23}$$

where each of the coefficients  $\mathbf{Y}_{m,k}(y)$  is an off-diagonal matrix, and where  $\mathbf{D}_m(y, \zeta)$  is a diagonal matrix with the asymptotic expansion

$$\mathbf{D}_m(y, \zeta) \sim \mathbf{D}_{m,-3}(y)\zeta^3 + \mathbf{D}_{m,-2}(y)\zeta^2 + \mathbf{D}_{m,-1}(y)\zeta + \mathbf{D}_{m,0}(y) \log(-\zeta) + \sum_{k=1}^{\infty} \mathbf{D}_{m,k}(y)\zeta^{-k}, \quad \zeta \rightarrow \infty, \tag{5.24}$$

where all of the coefficients  $\mathbf{D}_{m,k}(y)$  are diagonal matrices. The value of  $\lambda_m$  will be determined from the coefficient  $\mathbf{D}_{m,0}$ . By differentiation of the formal series, we obtain

$$\frac{\partial \mathbf{L}_m}{\partial \zeta} - \mathbf{V}_m \mathbf{L}_m \sim [\mathbf{M}_{m,-2} \zeta^2 + \mathbf{M}_{m,-1} \zeta + \mathbf{M}_{m,0} + \mathbf{M}_{m,1} \zeta^{-1} + \mathbf{M}_{m,2} \zeta^{-2} + \mathcal{O}(\zeta^{-3})] e^{\mathbf{D}_m(y, \zeta)}, \quad \zeta \rightarrow \infty, \quad (5.25)$$

where the matrix coefficients are systematically determined as follows:

$$\begin{aligned} \mathbf{M}_{m,-2} &:= 3\mathbf{D}_{m,-3}(y) + \frac{3}{2}\sigma_3 \\ \mathbf{M}_{m,-1} &:= 3\mathbf{Y}_{m,1}(y)\mathbf{D}_{m,-3}(y) + 2\mathbf{D}_{m,-2}(y) - 3 \begin{bmatrix} 0 & \mathcal{U}_m(y) \\ -\mathcal{V}_m(y) & 0 \end{bmatrix} + \frac{3}{2}\sigma_3 \mathbf{Y}_{m,1}(y) \\ \mathbf{M}_{m,0} &:= 3\mathbf{Y}_{m,2}(y)\mathbf{D}_{m,-3}(y) + 2\mathbf{Y}_{m,1}(y)\mathbf{D}_{m,-2}(y) + \mathbf{D}_{m,-1}(y) \\ &\quad - \frac{1}{2} \begin{bmatrix} -6\mathcal{U}_m(y)\mathcal{V}_m(y) - y & 2\mathcal{W}_m(y) \\ -2\mathcal{Z}_m(y) & 6\mathcal{U}_m(y)\mathcal{V}_m(y) + y \end{bmatrix} - 3 \begin{bmatrix} 0 & \mathcal{U}_m(y) \\ -\mathcal{V}_m(y) & 0 \end{bmatrix} \mathbf{Y}_{m,1}(y) + \frac{3}{2}\sigma_3 \mathbf{Y}_{m,2}(y) \\ \mathbf{M}_{m,1} &:= 3\mathbf{Y}_{m,3}(y)\mathbf{D}_{m,-3}(y) + 2\mathbf{Y}_{m,2}(y)\mathbf{D}_{m,-2}(y) + \mathbf{Y}_{m,1}(y)\mathbf{D}_{m,-1}(y) + \mathbf{D}_{m,0}(y) \\ &\quad - \frac{1}{2} \begin{bmatrix} -6\mathcal{U}_m(y)\mathcal{V}_m(y) - y & 2\mathcal{W}_m(y) \\ -2\mathcal{Z}_m(y) & 6\mathcal{U}_m(y)\mathcal{V}_m(y) + y \end{bmatrix} \mathbf{Y}_{m,1}(y) \\ &\quad - 3 \begin{bmatrix} 0 & \mathcal{U}_m(y) \\ -\mathcal{V}_m(y) & 0 \end{bmatrix} \mathbf{Y}_{m,2}(y) + \frac{3}{2}\sigma_3 \mathbf{Y}_{m,3}(y) \\ \mathbf{M}_{m,2} &:= -\mathbf{D}_{m,1}(y) - \frac{1}{2} \begin{bmatrix} 0 & 2\mathcal{W}_m(y) \\ -2\mathcal{Z}_m(y) & 0 \end{bmatrix} \mathbf{Y}_{m,2}(y) \\ &\quad - 3 \begin{bmatrix} 0 & \mathcal{U}_m(y) \\ -\mathcal{V}_m(y) & 0 \end{bmatrix} \mathbf{Y}_{m,3}(y) + \text{off-diagonal terms}, \end{aligned} \quad (5.26)$$

and so on. By setting each  $\mathbf{M}_{m,k}$  to zero in sequence, and by separating the diagonal and off-diagonal parts of each of these equations, we obtain the following: the matrix equation  $\mathbf{M}_{m,-2} = \mathbf{0}$  implies

$$\mathbf{D}_{m,-3}(y) = -\frac{1}{2}\sigma_3, \quad (5.27)$$

the matrix equation  $\mathbf{M}_{m,-1} = \mathbf{0}$  implies

$$\mathbf{Y}_{m,1}(y) = \begin{bmatrix} 0 & \mathcal{U}_m(y) \\ \mathcal{V}_m(y) & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{D}_{m,-2}(y) = \mathbf{0}, \quad (5.28)$$

the matrix equation  $\mathbf{M}_{m,0} = \mathbf{0}$  implies

$$\mathbf{Y}_{m,2}(y) = \frac{1}{3} \begin{bmatrix} 0 & \mathcal{W}_m(y) \\ \mathcal{Z}_m(y) & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{D}_{m,-1}(y) = -\frac{1}{2}y\sigma_3, \quad (5.29)$$

the matrix equation  $\mathbf{M}_{m,1} = \mathbf{0}$  implies

$$\mathbf{Y}_{m,3}(y) = -\frac{1}{3} \begin{bmatrix} 0 & 3\mathcal{U}_m(y)^2\mathcal{V}_m(y) + y\mathcal{U}_m(y) \\ 3\mathcal{U}_m(y)\mathcal{V}_m(y)^2 + y\mathcal{V}_m(y) & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{D}_{m,0}(y) = -3\lambda_m\sigma_3, \quad (5.30)$$

where we have recalled the definition (5.22) of  $\lambda_m$ , and finally, the diagonal terms of the matrix equation  $\mathbf{M}_{m,2} = \mathbf{0}$  imply that

$$\mathbf{D}_{m,1}(y) = -2H_m(y)\sigma_3 \quad (5.31)$$

where the Painlevé-II *Hamiltonian* is

$$H_m(y) := \frac{1}{6}\mathcal{W}_m(y)\mathcal{Z}_m(y) - \frac{3}{2}\mathcal{U}_m(y)^2\mathcal{V}_m(y)^2 - \frac{1}{2}y\mathcal{U}_m(y)\mathcal{V}_m(y). \quad (5.32)$$

It follows that the formal asymptotic expansion of  $\mathbf{L}_m(y, \zeta)$  for large  $\zeta$  takes the form

$$\begin{aligned} \mathbf{L}_m(y, \zeta) e^{(\zeta^3 + y\zeta)\sigma_3/2} (-\zeta)^{3\lambda_m\sigma_3} &\sim \mathbb{I} + [\mathbf{Y}_{m,1}(y) + \mathbf{D}_{m,1}(y)] \zeta^{-1} \\ &+ \left[ \mathbf{Y}_{m,2}(y) + \mathbf{D}_{m,2}(y) + \frac{1}{2} \mathbf{D}_{m,1}(y)^2 + \mathbf{Y}_{m,1}(y) \mathbf{D}_{m,1}(y) \right] \zeta^{-2} \\ &+ \mathcal{O}(\zeta^{-3}), \quad \zeta \rightarrow \infty. \end{aligned} \quad (5.33)$$

Comparing with (5.1) and (5.10), we see that

$$\lambda_m = \frac{1}{6} - \frac{m}{3}, \quad (5.34)$$

and also that

$$\begin{aligned} \mathbf{A}_m(y) &= \mathbf{Y}_{m,1}(y) + \mathbf{D}_{m,1}(y) \\ \mathbf{B}_m(y) &= \mathbf{Y}_{m,2}(y) + \mathbf{D}_{m,2}(y) + \frac{1}{2} \mathbf{D}_{m,1}(y)^2 + \mathbf{Y}_{m,1}(y) \mathbf{D}_{m,1}(y), \end{aligned} \quad (5.35)$$

so that in addition to the relations (5.16) we have

$$A_{m,11}(y) = -2H_m(y) \quad \text{and} \quad A_{m,22}(y) = 2H_m(y) \quad (5.36)$$

and

$$B_{m,12}(y) = \frac{1}{3} \mathcal{W}_m(y) + 2H_m(y) \mathcal{U}_m(y) \quad \text{and} \quad B_{m,21}(y) = \frac{1}{3} \mathcal{Z}_m(y) - 2H_m(y) \mathcal{V}_m(y). \quad (5.37)$$

Now that the value of  $\lambda_m$  has been determined according to (5.34), a direct calculation shows that the logarithmic derivatives

$$\mathcal{P}_m(y) := \frac{\mathcal{U}'_m(y)}{\mathcal{U}_m(y)} \quad \text{and} \quad \mathcal{Q}_m(y) := \frac{\mathcal{V}'_m(y)}{\mathcal{V}_m(y)} \quad (5.38)$$

satisfy the uncoupled Painlevé-II equations

$$\begin{aligned} \mathcal{P}_m''(y) &= 2\mathcal{P}_m(y)^3 + \frac{2}{3} y \mathcal{P}_m(y) - \left[ \frac{1}{3} - 2\lambda_m \right] & \mathcal{Q}_m''(y) &= 2\mathcal{Q}_m(y)^3 + \frac{2}{3} y \mathcal{Q}_m(y) - \left[ \frac{1}{3} + 2\lambda_m \right] \\ &= 2\mathcal{P}_m(y)^3 + \frac{2}{3} y \mathcal{P}_m(y) - \frac{2}{3} m & &= 2\mathcal{Q}_m(y)^3 + \frac{2}{3} y \mathcal{Q}_m(y) + \frac{2}{3} (m-1). \end{aligned} \quad (5.39)$$

Note that for general  $m \in \mathbb{Z}$ , both equations are of inhomogeneous type.

**5.3. Solution of Riemann-Hilbert Problem 4.1 for  $m = 0$ .** Recalling the relation (5.8) and the definitions (5.16) of the potentials  $\mathcal{U}_m$ ,  $\mathcal{V}_m$ ,  $\mathcal{W}_m$ , and  $\mathcal{Z}_m$ , we learn that when  $m = 0$ ,  $\mathcal{U}_0(y) \equiv \sigma$ , where (assuming for the moment continuity of  $\mathcal{U}_0(y)$  with respect to  $y$ )  $\sigma = \pm 1$  is a fixed sign to be determined. According to the system of differential equations (5.19) necessarily satisfied by the potentials, we can determine the values of the remaining potentials in the special case of  $m = 0$ :

$$\mathcal{U}_0(y) = \sigma, \quad \mathcal{V}_0(y) = -\frac{\sigma}{6} y, \quad \mathcal{W}_0(y) = 0, \quad \mathcal{Z}_0(y) = -\frac{\sigma}{2}. \quad (5.40)$$

Therefore, from (5.15) and (5.17) we see that in the special case of  $m = 0$ , the linear differential equations simultaneously satisfied by the matrix  $\mathbf{L}_0(a, \zeta) = \mathbf{Z}_0(\zeta; y) e^{-(\zeta^3 + y\zeta)\sigma_3/2}$  take the form

$$\frac{\partial \mathbf{L}_0}{\partial y} = \begin{bmatrix} -\frac{1}{2}\zeta & \sigma \\ \frac{1}{6}\sigma y & \frac{1}{2}\zeta \end{bmatrix} \mathbf{L}_0 \quad \text{and} \quad \frac{\partial \mathbf{L}_0}{\partial \zeta} = \begin{bmatrix} -\frac{3}{2}\zeta^2 & 3\sigma\zeta \\ \frac{1}{2}\sigma(y\zeta + 1) & \frac{3}{2}\zeta^2 \end{bmatrix} \mathbf{L}_0. \quad (5.41)$$

Let  $\mathbf{l} = (l_1, l_2)^\top$  denote either of the columns of  $\mathbf{L}_0$ . From the first of the above two Lax equations, we observe that

$$l_2 = \sigma \left( \frac{\partial l_1}{\partial y} + \frac{1}{2} \zeta l_1 \right) \quad (5.42)$$

and then that

$$l_1 = F(\xi), \quad \xi := \frac{1}{6^{1/3}} \left( y + \frac{3}{2} \zeta^2 \right), \quad (5.43)$$

where  $F$  is any solution of Airy's equation  $F''(\xi) = \xi F(\xi)$ . Therefore  $\mathbf{l}$  necessarily has the form

$$\mathbf{l} = \begin{bmatrix} F(\xi) \\ \sigma(6^{-1/3} F'(\xi) + \frac{1}{2} \zeta F(\xi)) \end{bmatrix} \quad (5.44)$$

and by substitution into the second equation of the Lax pair we learn that the Airy function  $F$  may only depend on  $\zeta$  through the variable  $\xi$ , in other words, we must select for  $F$  a linear combination of  $\text{Ai}(\xi)$  and  $\text{Bi}(\xi)$  with coefficients independent of  $\zeta$ .

Now we try to select the correct solutions of Airy's equation to construct  $\mathbf{L}_0(y, \zeta)$  and hence  $\mathbf{Z}_0(\zeta; y)$  in each of the six sectors of analyticity in such a way that we obtain the desired asymptotics for large  $\zeta$  and satisfy the required jump conditions. This procedure will also determine the correct value of  $\sigma = \pm 1$ .

Let us consider three specific choices of solution to Airy's equation leading to the three solution vectors

$$\mathbf{l}^0 := \begin{bmatrix} \text{Ai}(\xi) \\ \sigma(6^{-1/3} \text{Ai}'(\xi) + \frac{1}{2} \zeta \text{Ai}(\xi)) \end{bmatrix}, \quad \mathbf{l}^\pm := \begin{bmatrix} \text{Ai}(e^{\pm 2\pi i/3} \xi) \\ \sigma(6^{-1/3} e^{\pm 2\pi i/3} \text{Ai}'(e^{\pm 2\pi i/3} \xi) + \frac{1}{2} \zeta \text{Ai}(e^{\pm 2\pi i/3} \xi)) \end{bmatrix}. \quad (5.45)$$

Each of these is an entire function of  $\zeta$ . From the standard asymptotic formulae

$$\text{Ai}(z) = \frac{1}{2\sqrt{\pi}} z^{-1/4} e^{-2z^{3/2}/3} (1 + \mathcal{O}(|z|^{-3/2})) \quad \text{and} \quad \text{Ai}'(z) = -\frac{1}{2\sqrt{\pi}} z^{1/4} e^{-2z^{3/2}/3} (1 + \mathcal{O}(|z|^{-3/2})) \quad (5.46)$$

valid as  $z \rightarrow \infty$  with  $|\arg(z)| \leq \pi - \delta$  for any  $\delta > 0$ , we see that (all of the following asymptotic statements assume  $y$  is held fixed)

$$\mathbf{l}^0 = \frac{\zeta^{-1/2} e^{-(\zeta^3 + y\zeta)/2}}{(48)^{1/6} \sqrt{\pi}} \begin{bmatrix} 1 + \mathcal{O}(\zeta^{-1}) \\ \mathcal{O}(\zeta^{-1}) \end{bmatrix}, \quad \zeta \rightarrow \infty, \quad |\arg(\zeta)| \leq \frac{\pi}{2} - \delta, \quad (5.47)$$

$$\mathbf{l}^0 = -\sigma \frac{(-\zeta)^{1/2} e^{(\zeta^3 + y\zeta)/2}}{(48)^{1/6} \sqrt{\pi}} \begin{bmatrix} \mathcal{O}(\zeta^{-1}) \\ 1 + \mathcal{O}(\zeta^{-1}) \end{bmatrix}, \quad \zeta \rightarrow \infty, \quad |\arg(-\zeta)| \leq \frac{\pi}{2} - \delta. \quad (5.48)$$

Similarly,

$$\mathbf{l}^\pm = -\sigma e^{\pm 2\pi i/3} \frac{(e^{\pm i\pi/3} \zeta)^{1/2} e^{(\zeta^3 + y\zeta)/2}}{(48)^{1/6} \sqrt{\pi}} \begin{bmatrix} \mathcal{O}(\zeta^{-1}) \\ 1 + \mathcal{O}(\zeta^{-1}) \end{bmatrix}, \quad \zeta \rightarrow \infty, \quad |\arg(e^{\pm i\pi/3} \zeta)| \leq \frac{\pi}{2} - \delta, \quad (5.49)$$

$$\mathbf{l}^\pm = \frac{(e^{\mp 2\pi i/3} \zeta)^{-1/2} e^{-(\zeta^3 + y\zeta)/2}}{(48)^{1/6} \sqrt{\pi}} \begin{bmatrix} 1 + \mathcal{O}(\zeta^{-1}) \\ \mathcal{O}(\zeta^{-1}) \end{bmatrix}, \quad \zeta \rightarrow \infty, \quad |\arg(e^{\mp 2\pi i/3} \zeta)| \leq \frac{\pi}{2} - \delta. \quad (5.50)$$

We may now define a candidate for  $\mathbf{Z}_0(\zeta; y)$  that satisfies the required analyticity and normalization properties among the conditions of Riemann-Hilbert Problem 4.1:

$$\mathbf{Z}_0(\zeta; y) := \begin{cases} (48)^{1/6} \sqrt{\pi} (i \mathbf{l}^0, \sigma e^{-2\pi i/3} \mathbf{l}^-) e^{(\zeta^3 + y\zeta)\sigma_3/2}, & \arg(\zeta) \in (0, \frac{\pi}{3}) \\ (48)^{1/6} \sqrt{\pi} (e^{i\pi/6} \mathbf{l}^+, \sigma e^{-2\pi i/3} \mathbf{l}^-) e^{(\zeta^3 + y\zeta)\sigma_3/2}, & \arg(\zeta) \in (\frac{\pi}{3}, \frac{2\pi}{3}) \\ (48)^{1/6} \sqrt{\pi} (e^{i\pi/6} \mathbf{l}^+, -\sigma \mathbf{l}^0) e^{(\zeta^3 + y\zeta)\sigma_3/2}, & \arg(\zeta) \in (\frac{2\pi}{3}, \pi) \\ (48)^{1/6} \sqrt{\pi} (e^{-i\pi/6} \mathbf{l}^-, -\sigma \mathbf{l}^0) e^{(\zeta^3 + y\zeta)\sigma_3/2}, & \arg(\zeta) \in (-\pi, -\frac{2\pi}{3}) \\ (48)^{1/6} \sqrt{\pi} (e^{-i\pi/6} \mathbf{l}^-, \sigma e^{2\pi i/3} \mathbf{l}^+) e^{(\zeta^3 + y\zeta)\sigma_3/2}, & \arg(\zeta) \in (-\frac{2\pi}{3}, -\frac{\pi}{3}) \\ (48)^{1/6} \sqrt{\pi} (-i \mathbf{l}^0, \sigma e^{2\pi i/3} \mathbf{l}^+) e^{(\zeta^3 + y\zeta)\sigma_3/2}, & \arg(\zeta) \in (-\frac{\pi}{3}, 0). \end{cases} \quad (5.51)$$

It only remains to verify the jump conditions of Riemann-Hilbert Problem 4.1, with the help of the identity

$$\text{Ai}(z) + e^{-2\pi i/3} \text{Ai}(ze^{-2\pi i/3}) + e^{2\pi i/3} \text{Ai}(ze^{2\pi i/3}) = 0, \quad z \in \mathbb{C}, \quad (5.52)$$

which implies the vector identity  $\mathbf{l}^0 + e^{-2\pi i/3} \mathbf{l}^- + e^{2\pi i/3} \mathbf{l}^+ = \mathbf{0}$ . Using this fact, it is a direct matter to confirm that the required jump conditions are indeed satisfied, *provided we make the choice of sign  $\sigma = +1$* . We have therefore proved the following.

**Proposition 5.1.** *Let  $m = 0$ . For every  $y \in \mathbb{R}$ , Riemann-Hilbert Problem 4.1 has a unique solution given explicitly by the formulae (5.45) and (5.51) with  $\sigma = +1$ . Moreover,  $\mathbf{Z}_0(\zeta; y)$  is an entire function of  $y$ ,  $\mathbf{Z}_0(\zeta; y)(-\zeta)^{\sigma_3/2}$  has a complete asymptotic expansion as  $\zeta \rightarrow \infty$  in descending integer powers of  $\zeta$  (as in (5.1)) differentiable term-by-term with respect to both  $y$  and  $\zeta$ , and the corresponding potentials  $\mathcal{U}_0(y)$ ,  $\mathcal{V}_0(y)$ ,  $\mathcal{W}_0(y)$  and  $\mathcal{Z}_0(y)$  are given by (5.40) with  $\sigma = +1$ .*

**5.4. Schlesinger-Bäcklund transformations. Solution of Riemann-Hilbert Problem 4.1 for general  $m \in \mathbb{Z}$ .** Now we describe an inductive procedure for obtaining  $\mathbf{Z}_m(\zeta; y)$  uniquely from the conditions of Riemann-Hilbert Problem 4.1 for  $m \neq 0$ . Suppose  $y \in \mathbb{R}$  is a value for which Riemann-Hilbert Problem 4.1 has a (unique) solution  $\mathbf{Z}_m(\zeta; y)$  for some  $m \in \mathbb{Z}$ . Following [18, Chapter 6], we claim that

$$\mathbf{Z}_{m\pm 1}(\zeta; y) = (\mathbf{S}_{m,1}^\pm(y)\zeta + \mathbf{S}_{m,0}^\pm(y))\mathbf{Z}_m(\zeta; y) \quad (5.53)$$

if the matrices  $\mathbf{S}_{m,0}^\pm(y)$  and  $\mathbf{S}_{m,1}^\pm(y)$  can be properly chosen.

It is immediately clear that since the prefactor is entire in  $\zeta$ , the right-hand side of (5.53) has the same domain of analyticity, achieves its boundary values in the same classical sense, and satisfies exactly the same jump conditions, as does  $\mathbf{Z}_m(\zeta; y)$ . Of course the analyticity and jump conditions in Riemann-Hilbert Problem 4.1 are independent of  $m$ , so it only remains to impose the normalization condition (4.64) on  $\mathbf{Z}_{m\pm 1}(\zeta; y)$  given by (5.53), and this condition takes the form

$$\lim_{\zeta \rightarrow \infty} (\mathbf{S}_{m,1}^\pm(y)\zeta + \mathbf{S}_{m,0}^\pm(y))\mathbf{Z}_m(\zeta; y)(-\zeta)^{(1-2m\mp 2)\sigma_3/2} = \mathbb{I}. \quad (5.54)$$

Assuming the expansion (5.1) (which at the moment is justified only for  $m = 0$ ), we are requiring that

$$(\mathbf{S}_{m,1}^\pm(y)\zeta + \mathbf{S}_{m,0}^\pm(y))(\mathbb{I} + \mathbf{A}_m(y)\zeta^{-1} + \mathbf{B}_m(y)\zeta^{-2} + \mathbf{C}_m(y)\zeta^{-3} + \mathcal{O}(\zeta^{-4}))(-\zeta)^{\mp\sigma_3} = \mathbb{I} + \mathcal{O}(\zeta^{-1}), \quad \zeta \rightarrow \infty. \quad (5.55)$$

Equating the polynomial part of the left-hand side to the identity yields the conditions

$$\begin{aligned} \mathbf{S}_{m,1}^+(y) \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} &= \mathbf{0} \\ (\mathbf{S}_{m,1}^+(y)\mathbf{A}_m(y) + \mathbf{S}_{m,0}^+(y)) \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} &= \mathbf{0} \\ (\mathbf{S}_{m,1}^+(y)\mathbf{B}_m(y) + \mathbf{S}_{m,0}^+(y)\mathbf{A}_m(y)) \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} + \mathbf{S}_{m,1}^+(y) \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} &= -\mathbb{I} \end{aligned} \quad (5.56)$$

which can be solved uniquely for the matrix elements of  $\mathbf{S}_{m,1}^+(y)$  and  $\mathbf{S}_{m,0}^+(y)$ :

$$\begin{aligned} \mathbf{S}_{m,1}^+(y) &= \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}, \\ \mathbf{S}_{m,0}^+(y) &= \begin{bmatrix} B_{m,12}(y)/A_{m,12}(y) - A_{m,22}(y) & A_{m,12}(y) \\ -1/A_{m,12}(y) & 0 \end{bmatrix} = \begin{bmatrix} \mathcal{W}_m(y)/(3\mathcal{U}_m(y)) & \mathcal{U}_m(y) \\ -1/\mathcal{U}_m(y) & 0 \end{bmatrix}, \end{aligned} \quad (5.57)$$

provided that  $\mathcal{U}_m(y) \neq 0$ , where we have used (5.16), (5.36), and (5.37). Similarly, from (5.55) we obtain the conditions

$$\begin{aligned} \mathbf{S}_{m,1}^-(y) \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} &= \mathbf{0} \\ (\mathbf{S}_{m,1}^-(y)\mathbf{A}_m(y) + \mathbf{S}_{m,0}^-(y)) \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} &= \mathbf{0} \\ (\mathbf{S}_{m,1}^-(y)\mathbf{B}_m(y) + \mathbf{S}_{m,0}^-(y)\mathbf{A}_m(y)) \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \mathbf{S}_{m,1}^-(y) \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} &= -\mathbb{I} \end{aligned} \quad (5.58)$$

which can be solved uniquely for the matrix elements of  $\mathbf{S}_{m,1}^-(y)$  and  $\mathbf{S}_{m,0}^-(y)$ :

$$\begin{aligned} \mathbf{S}_{m,1}^-(y) &= \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}, \\ \mathbf{S}_{m,0}^-(y) &= \begin{bmatrix} 0 & -1/A_{m,21}(y) \\ A_{m,21}(y) & B_{m,21}(y)/A_{m,21}(y) - A_{m,11}(y) \end{bmatrix} = \begin{bmatrix} 0 & -1/\mathcal{V}_m(y) \\ \mathcal{V}_m(y) & \mathcal{Z}_m(y)/(3\mathcal{V}_m(y)) \end{bmatrix}, \end{aligned} \quad (5.59)$$

provided that  $\mathcal{V}_m(y) \neq 0$ . When it makes sense, the formula (5.53) subject to (5.57) or (5.59) constitutes a so-called *discrete isomonodromic Schlesinger transformation*.

The recurrence relation (5.53) also implies a corresponding recurrence for the potentials  $\mathcal{U}_m$ ,  $\mathcal{V}_m$ ,  $\mathcal{W}_m$ , and  $\mathcal{Z}_m$ . Indeed, expanding the right-hand side of (5.55) as  $\mathbb{I} + \mathbf{A}_{m\pm 1}(y)\zeta^{-1} + \mathcal{O}(\zeta^{-2})$  the terms proportional

to  $\zeta^{-1}$  yield the conditions

$$(\mathbf{S}_{m,1}^+(y)\mathbf{C}_m(y) + \mathbf{S}_{m,0}^+(y)\mathbf{B}_m(y)) \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} + (\mathbf{S}_{m,1}^+(y)\mathbf{A}_m(y) + \mathbf{S}_{m,0}^+(y)) \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = -\mathbf{A}_{m+1}(y) \quad (5.60)$$

and

$$(\mathbf{S}_{m,1}^-(y)\mathbf{C}_m(y) + \mathbf{S}_{m,0}^-(y)\mathbf{B}_m(y)) \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + (\mathbf{S}_{m,1}^-(y)\mathbf{A}_m(y) + \mathbf{S}_{m,0}^-(y)) \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = -\mathbf{A}_{m-1}(y). \quad (5.61)$$

Recalling the definition (5.16) of the potentials  $(\mathcal{U}_m, \mathcal{V}_m)$  in terms of the matrix entries of  $\mathbf{A}_m(y)$ , we examine the  $(2, 1)$ -entry of the matrix equation (5.60) and easily obtain the relation  $\mathcal{V}_{m+1}(y) = 1/\mathcal{U}_m(y)$ . Then, from the differential equations (5.19) satisfied by the potentials for all  $n$ , we obtain

$$\mathcal{U}_{m+1}(y) = -\frac{1}{6}y\mathcal{U}_m(y) - \frac{\mathcal{U}_m'(y)^2}{\mathcal{U}_m(y)} + \frac{1}{2}\mathcal{U}_m''(y) \quad \text{and} \quad \mathcal{V}_{m+1}(y) = \frac{1}{\mathcal{U}_m(y)} \quad (5.62)$$

which together with the implied relations for  $\mathcal{W}_{m+1}$  and  $\mathcal{Z}_{m+1}$  constitutes a *Bäcklund transformation* for the Painlevé-II system (5.19). Similarly, the  $(1, 2)$ -element of (5.61) together with (5.19) yields the Bäcklund transformation

$$\mathcal{U}_{m-1}(y) = \frac{1}{\mathcal{V}_m(y)} \quad \text{and} \quad \mathcal{V}_{m-1}(y) = \frac{1}{2}\mathcal{V}_m''(y) - \frac{\mathcal{V}_m'(y)^2}{\mathcal{V}_m(y)} - \frac{1}{6}y\mathcal{V}_m(y), \quad (5.63)$$

which is easily seen to be inverse to (5.62). The iterative scheme for generating the solution to Riemann-Hilbert Problem 4.1 for general  $m \in \mathbb{Z}$  is then simply to start with the solution valid for all  $y \in \mathbb{R}$  for  $m = 0$  as described by Proposition 5.1 and attempt to apply the discrete isomonodromic Schlesinger transformation (5.53) in order to repeatedly increase or decrease the value of  $m$  in integer steps. At the level of the implied Bäcklund transformations (5.62) and (5.63), this procedure generates a family  $\{(\mathcal{U}_m(y), \mathcal{V}_m(y), \mathcal{W}_m(y), \mathcal{Z}_m(y))\}_{m \in \mathbb{Z}}$  of *rational solutions* of the Painlevé-II system (5.19). At the level of the logarithmic derivatives given by (5.38) we have a family of rational solutions of inhomogeneous Painlevé-II equations with certain quantized values for the inhomogeneity parameter  $\alpha$ . In fact, it is known [22] that the equation

$$\mathcal{P}''(y) = 2\mathcal{P}(y)^3 + \frac{2}{3}y\mathcal{P}(y) - \frac{2}{3}\alpha \quad (5.64)$$

has a unique rational solution  $\mathcal{P} = \mathcal{P}_\alpha(y)$  if and only if  $\alpha \in \mathbb{Z}$ . These may be represented as logarithmic derivatives of other rational functions  $\mathcal{U}_\alpha$  that are ratios of consecutive *Yablonskii-Vorob'ev polynomials* [13, 14], which are known to have a number of remarkable properties.

It is important for us to recognize that the poles of these solutions can occur for finite real values of  $y$ . For example, applying (5.62) twice with  $\mathcal{U}_0(y) = 1$  we obtain  $\mathcal{U}_1(y) = -y/6$  and  $\mathcal{U}_2(y) = (y^3 + 6)/(36y)$ , and the latter has a pole at  $y = 0$ . These poles correspond to values of  $y$  for which the discrete Schlesinger maps from the neighboring values of  $m$  fail to exist, and consequently for these values of  $y$  and  $m$  *there exists no solution to Riemann-Hilbert Problem 4.1*.

While there exist values of  $y$  for which the Bäcklund transformations (5.62) or (5.63) may not make sense, we may always interpret these transformations as maps on the ring of rational functions, and as such they can be shown to have a kind of *singularity confinement property*: if  $y_0 \in \mathbb{C}$  is a pole of the solution for some  $m$ , then  $y_0$  is a regular point of the solution for  $m \pm 1$ . Indeed, a local analysis of the system (5.20) in the spirit of Painlevé's original method [19] shows that all poles of  $(\mathcal{U}_m(y), \mathcal{V}_m(y))$  are necessarily simple and simultaneously occur in both functions; if  $y_0$  is a pole, then for some constants  $k \in \mathbb{C} \setminus \{0\}$  and  $\omega \in \mathbb{C}$  we necessarily have

$$\begin{aligned} \mathcal{U}_m(y) &= k \left[ \frac{1}{y - y_0} + \frac{y_0}{18}(y - y_0) + \left( \frac{1}{12} + \omega \right) (y - y_0)^2 + U_3(y) \right] \\ \mathcal{V}_m(y) &= -k^{-1} \left[ \frac{1}{y - y_0} + \frac{y_0}{18}(y - y_0) + \left( \frac{1}{12} - \omega \right) (y - y_0)^2 + V_3(y) \right], \end{aligned} \quad (5.65)$$

where  $U_3(y)$  and  $V_3(y)$  are analytic functions each vanishing to third order at  $y = y_0$ . Then, using these formulae in the Bäcklund transformations (5.62)–(5.63) shows that the singularities at  $y = y_0$  in the pairs  $(\mathcal{U}_{m+1}(y), \mathcal{V}_{m+1}(y))$  and  $(\mathcal{U}_{m-1}(y), \mathcal{V}_{m-1}(y))$  are all removable. Moreover, both  $\mathcal{V}_{m+1}(y)$  and  $\mathcal{U}_{m-1}(y)$  have simple zeros at  $y = y_0$ . It is easy to see that the singularity confinement property extends to the Schlesinger

transformations that generated (5.62)–(5.63). That is, if  $\mathbf{Z}_m(\zeta; y_0)$  does not exist (because  $y_0$  is a pole of  $\mathcal{U}_m(y)$  and  $\mathcal{V}_m(y)$ ), then both  $\mathbf{Z}_{m+1}(\zeta; y_0)$  and  $\mathbf{Z}_{m-1}(\zeta; y_0)$  do exist and they may be calculated from  $\mathbf{Z}_m(\zeta; y)$  by applying Schlesinger transformations and then taking the limit  $y \rightarrow y_0$ . We summarize these results in the following proposition.

**Proposition 5.2.** *Let  $m \in \mathbb{Z}$ . Then by iterated Schlesinger transformations given by (5.53) subject to (5.57) or (5.59) applied to the base case of  $m = 0$  characterized by Proposition 5.1, a matrix function  $\mathbf{Z}_m(\zeta; y)$  is well-defined as a rational function of  $y$  with simple poles. The poles of  $\mathbf{Z}_m(\zeta; y)$  are exactly the poles of the corresponding potentials  $(\mathcal{U}_m(y), \mathcal{V}_m(y))$  obtained from the base case of  $(\mathcal{U}_0(y), \mathcal{V}_0(y)) = (1, -y/6)$  by iterated Bäcklund transformations (5.62)–(5.63).*

- If  $y_0$  is not a pole, then there is a neighborhood  $A \subset \mathbb{C}$  with  $y_0 \in A$  such that for  $y \in A$ , the matrix  $\mathbf{Z}_m(\zeta; y)$ 
  - is analytic in  $y$ ,
  - is the solution of Riemann-Hilbert Problem 4.1,
  - has an expansion for large  $\zeta$  of the form (5.1) differentiable with respect to both  $y$  and  $\zeta$  and uniform for  $y \in A$ , and
  - is uniformly bounded for  $y \in A$  and  $\zeta$  in compact subsets of  $\mathbb{C}$ .
- If  $y_0$  is a pole, then Riemann-Hilbert Problem 4.1 has no solution for  $y = y_0$ ; however both  $\mathbf{Z}_{m+1}(\zeta; y)$  and  $\mathbf{Z}_{m-1}(\zeta; y)$  are regular at  $y_0$  (in the sense of taking limits from  $y \neq y_0$ ), and both  $\mathcal{V}_{m+1}(y)$  and  $\mathcal{U}_{m-1}(y)$  have simple zeros at  $y = y_0$ .

Since the poles are simple, the bound on the elements of  $\mathbf{Z}_m(\zeta; y)$  (and its inverse, since  $\det(\mathbf{Z}_m(\zeta; y)) = 1$  where defined) can be strengthened to  $\mathcal{O}(|y - \mathcal{P}(\mathcal{U}_m)|^{-1})$  (or equivalently  $\mathcal{O}(|y - \mathcal{P}(\mathcal{V}_m)|^{-1})$ ) uniformly on compact sets in the  $\zeta$ -plane. This improved bound also applies to all of the coefficients in the expansion (5.1) and to the error term as well. In particular, this implies that at  $y = y_0 \in \mathcal{P}(\mathcal{U}_m) = \mathcal{P}(\mathcal{V}_m)$ , the functions  $H_m(y) = \frac{1}{2}A_{m,22}(y) = -\frac{1}{2}A_{m,11}(y)$ ,  $B_{m,12}(y) = 2H_m(y)\mathcal{U}_m(y) - \mathcal{U}'_m(y)$ , and  $B_{m,21}(y) = \mathcal{V}'_m(y) - 2H_m(y)\mathcal{V}_m(y)$  all have simple poles.

## 6. INADEQUACY OF THE GLOBAL MODEL

On one hand, it seems quite reasonable to expect that the global model  $\dot{\mathbf{P}}_m(w)$  defined by (4.66) should provide a good approximation to  $\mathbf{P}(w)$ , at least if the integer  $m$  is properly chosen. Indeed, we have shown in an earlier paper [7] that far from criticality a global model obtained by steps completely analogous to those we have followed here (choice of an appropriate  $g$ -function, opening of lenses, pointwise approximation of jump matrices away from points of nonuniformity and the solution of a corresponding outer model problem, and exact solution of the jump conditions near points of nonuniformity yielding inner models that match well onto the outer model at disk boundaries) leads directly to a Riemann-Hilbert problem for the error of small-norm type. On the other hand, in the present case the formally large exponents  $\pm s/\epsilon_N$  were removed from the jump matrices in our construction of the inner model, and we may expect them to reappear when the mismatch between the inner and outer models is calculated on  $\partial U$ , contaminating the error estimates when  $s$  is small but large compared with  $\epsilon_N$ .

In order to properly gauge our prospects for success, we introduce a new matrix unknown  $\mathbf{Q}_m(w)$  constructed from the old unknown  $\mathbf{P}(w)$  and the explicit global model  $\dot{\mathbf{P}}_m(w)$  defined by (4.66) as follows:

$$\mathbf{Q}_m(w) := \mathbf{P}(w)\dot{\mathbf{P}}_m(w)^{-1}. \quad (6.1)$$

We might expect to be able to prove that  $\mathbf{Q}_m(w)$  is a small perturbation of the identity matrix with the use of small-norm theory applied to the Riemann-Hilbert problem satisfied by this matrix function.

To begin to determine the nature of the Riemann-Hilbert problem satisfied by  $\mathbf{Q}_m(w)$ , firstly note that  $\mathbf{Q}_m(w)$  is analytic for  $w \in U$ , since for such  $w$ ,  $\dot{\mathbf{P}}_m(w) = \dot{\mathbf{P}}_m^{\text{in}}(w)$  has determinant one and satisfies exactly the same jump conditions as does  $\mathbf{P}(w)$ . For  $w \in \mathbb{C} \setminus \bar{U}$ ,  $\mathbf{Q}_m(w)$  is analytic except on the arcs of the jump contour for  $\mathbf{P}(w)$  illustrated in Figure 3.2; it can further be checked that since both  $\mathbf{P}(w)$  and  $\dot{\mathbf{P}}_m(w) = \dot{\mathbf{P}}_m^{\text{out}}(w)$  change sign across the contour segment  $(\mathfrak{b}, 0)$ ,  $\mathbf{Q}_m(w)$  also extends analytically to this segment. The jump contour for  $\mathbf{Q}_m(w)$  also contains  $\partial U$ , where  $\mathbf{P}(w)$  is analytic but  $\dot{\mathbf{P}}_m(w)$  has a jump discontinuity stemming from the mismatch (because  $\zeta$  is large but finite on  $\partial U$ ) between  $\dot{\mathbf{P}}_m^{\text{in}}(w)$  and  $\dot{\mathbf{P}}_m^{\text{out}}(w)$ . The jump contour for  $\mathbf{Q}_m(w)$  is illustrated in Figure 6.1.

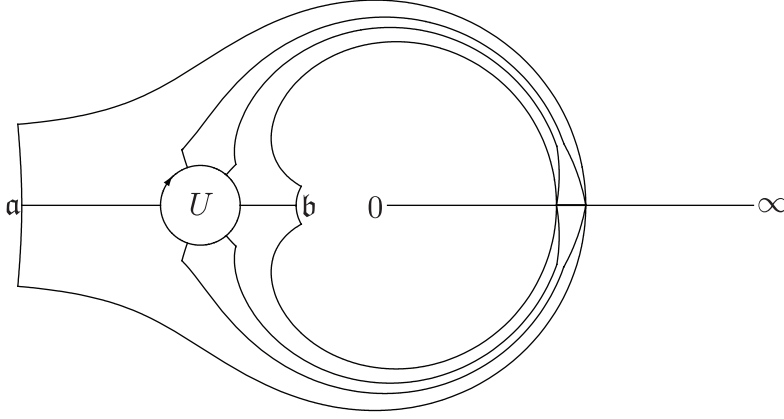


FIGURE 6.1. The contour  $\Sigma$  of discontinuity of the sectionally analytic function  $\mathbf{Q}_m(w)$ . With the exception of the circle  $\partial U$  which is clockwise-oriented as shown, all other arcs are oriented exactly as components of the jump contour of  $\mathbf{P}(w)$ .

Let  $\Sigma$  denote the jump contour for  $\mathbf{Q}_m(w)$ , that is,  $\mathbf{Q}_m(w)$  is analytic for  $w \in \mathbb{C} \setminus \Sigma$ , and it takes its boundary values in the classical sense on this contour from each component of the domain of analyticity. Also, since both  $\mathbf{P}(w)$  and  $\dot{\mathbf{P}}_m(w)$  tend to the identity matrix as  $w \rightarrow \infty$  (the former by hypothesis and the latter by the explicit formula for  $\dot{\mathbf{P}}_m^{\text{out}}(w)$ ), we have  $\mathbf{Q}_m(w) \rightarrow \mathbb{I}$  as  $w \rightarrow \infty$ .

For fixed  $m \in \mathbb{Z}$ , it turns out that for  $(x, t)$  sufficiently close to criticality,  $\mathbf{Q}_{m+}(\xi) = \mathbf{Q}_{m-}(\xi)(\mathbb{I} + \mathcal{O}(\epsilon_N))$  as  $\epsilon_N \rightarrow 0$  for  $\xi \in \Sigma \setminus (\mathbb{R}_+ \cup \partial U)$ , with the estimate being uniform with respect to  $\xi$  on the specified contour arcs and  $(x, t)$  near criticality. Indeed, exactly at criticality we have  $\mathbf{g}(w) = g(w)$  where  $g(w) = g(w; x, t)$  is as defined in [7, Section 4] and so by continuity with respect to  $x$  and  $t$  we will have near criticality that the jump matrix for  $\mathbf{O}(w)$  is subject to the same error estimates as recorded in [7, Proposition 5.1] generally valid away from criticality, where the neighborhoods  $U_1$  and  $U_2$  in the statement of the proposition are taken to coincide with  $U$ . Making the transformation from  $\mathbf{O}(w)$  to  $\mathbf{P}(w)$  we learn that the jump condition for  $\mathbf{P}(w)$  is of the form  $\mathbf{P}_+(\xi) = \mathbf{P}_-(\xi)(\mathbb{I} + \mathcal{O}(\epsilon_N))$  uniformly for all  $\xi$  in the jump contour pictured in Figure 3.2 with the exception of  $\xi \in U$  and on the ray  $\xi > -1$ . But,  $\dot{\mathbf{P}}_m(w)$  satisfies exactly the same jump conditions as does  $\mathbf{P}(w)$  within  $U$  and in the interval  $-1 < \xi < 0$ . The estimate on the jump conditions for  $\mathbf{Q}_m(w)$  asserted at the beginning of this paragraph then follow by exact computation.

For  $\xi \in \mathbb{R}_+ \setminus I$  (recall that  $I$  is a small open interval containing  $\xi = 1$  as shown in Figure 1.6), both  $\mathbf{P}(w)$  and  $\dot{\mathbf{P}}_m(w) = \dot{\mathbf{P}}_m^{\text{out}}(w)$  satisfy the same jump condition, from which it follows that for  $\xi > 0$  but  $\xi \notin I$ ,

$$\mathbf{Q}_{m+}(\xi) = \mathbf{P}_+(\xi)\dot{\mathbf{P}}_{m+}^{\text{out}}(\xi)^{-1} = \sigma_2 \mathbf{P}_-(\xi) \sigma_2 \left[ \sigma_2 \dot{\mathbf{P}}_{m-}^{\text{out}}(\xi) \sigma_2 \right]^{-1} = \sigma_2 \mathbf{P}_-(\xi) \dot{\mathbf{P}}_{m-}^{\text{out}}(\xi)^{-1} \sigma_2 = \sigma_2 \mathbf{Q}_{m-}(\xi) \sigma_2. \quad (6.2)$$

For  $\xi \in I$ , it follows from the fact that  $\mathbf{g}_+(\xi) + \mathbf{g}_-(\xi) = 0$  for  $\xi > 0$  that [7, Proposition 3.3] holds, implying that  $\mathbf{P}(w)$  satisfies  $\mathbf{P}_+(\xi) = \sigma_2 \mathbf{P}_-(\xi) \sigma_2 (\mathbb{I} + \mathcal{O}(\epsilon_N))$ . On the other hand, by construction  $\dot{\mathbf{P}}_m(w)$  satisfies the corresponding jump condition without error term. Because  $\dot{\mathbf{P}}_m(w)$  is uniformly bounded with bounded inverse, by a calculation similar to (6.2) we see that  $\mathbf{Q}_{m+}(\xi) = \sigma_2 \mathbf{Q}_{m-}(\xi) \sigma_2 (\mathbb{I} + \mathcal{O}(\epsilon_N))$  holds uniformly for  $\xi \in I \subset \mathbb{R}_+$ .

To compute the jump condition for  $\mathbf{Q}_m(w)$  across the negatively-oriented boundary  $\partial U$  of the disk  $U$ , first use the fact that  $\mathbf{P}(w)$  is continuous across  $\partial U$  to obtain

$$\mathbf{Q}_{m+}(\xi) = \mathbf{P}(\xi) \dot{\mathbf{P}}_m^{\text{out}}(\xi)^{-1} = \mathbf{P}(\xi) \dot{\mathbf{P}}_m^{\text{in}}(\xi)^{-1} \left[ \dot{\mathbf{P}}_m^{\text{in}}(\xi) \dot{\mathbf{P}}_m^{\text{out}}(\xi)^{-1} \right] = \mathbf{Q}_{m-}(\xi) \left[ \dot{\mathbf{P}}_m^{\text{in}}(\xi) \dot{\mathbf{P}}_m^{\text{out}}(\xi)^{-1} \right], \quad \xi \in \partial U. \quad (6.3)$$

By substitution from (4.60) and (4.65), we see that

$$\begin{aligned} & \dot{\mathbf{P}}_m^{\text{in}}(\xi) \dot{\mathbf{P}}_m^{\text{out}}(\xi)^{-1} \\ &= \epsilon_N^{(2m-1)\sigma_3/6} e^{\frac{1}{2}s\sigma_3/\epsilon_N} \left[ \eta_m(\xi)^{\sigma_3} \mathbf{Z}_m(\zeta; y) (-\zeta)^{(1-2m)\sigma_3/2} \eta_m(\xi)^{-\sigma_3} \right] e^{-\frac{1}{2}s\sigma_3/\epsilon_N} \epsilon_N^{(1-2m)\sigma_3/6} d_N(\xi)^{-\sigma_3}, \end{aligned} \quad (6.4)$$



for  $\xi \in \partial U$ , where  $\zeta = \epsilon_N^{-1/3} W(\xi)$ .

We therefore see that for  $(x, t)$  close enough to criticality,  $\mathbf{Q}_m(w)$  satisfies the conditions of the following type of Riemann-Hilbert problem.

**Riemann-Hilbert Problem 6.1.** *Seek a matrix  $\mathbf{Q}_m(w)$  with the following properties:*

**Analyticity:**  $\mathbf{Q}_m(w)$  is analytic for  $w \in \mathbb{C} \setminus \Sigma$ , where  $\Sigma$  is the contour (independent of  $x, t, m$ , and  $\epsilon_N$ ) pictured in Figure 6.1, and in each component of the domain of analyticity is uniformly Hölder continuous up to the boundary.

**Jump conditions:** The boundary values taken by  $\mathbf{Q}_m(w)$  on  $\Sigma$  are related as follows. For  $\xi \in \partial U \subset \Sigma$  with clockwise orientation,

$$\mathbf{Q}_{m+}(\xi) = \mathbf{Q}_{m-}(\xi) \mathbf{V}_{\mathbf{Q}_m}(\xi) \quad (6.5)$$

with jump matrix  $\mathbf{V}_{\mathbf{Q}_m}(\xi)$  defined by the right-hand side of (6.4). For  $\xi \in \mathbb{R}_+$ ,

$$\mathbf{Q}_{m+}(\xi) = \sigma_2 \mathbf{Q}_{m-}(\xi) \sigma_2 (\mathbb{I} + \mathbf{X}_m(\xi)) \quad (6.6)$$

where  $\mathbf{X}_m(\xi) = \mathcal{O}(\epsilon_N)$  holds uniformly for  $\xi \in I$  and  $\mathbf{X}_m(\xi) = 0$  for  $\xi \in \mathbb{R}_+ \setminus I$ . For all remaining  $\xi \in \Sigma$  we have the uniform estimate

$$\mathbf{Q}_{m+}(\xi) = \mathbf{Q}_{m-}(\xi) (\mathbb{I} + \mathcal{O}(\epsilon_N)) \quad (6.7)$$

as  $\epsilon_N \downarrow 0$ .

**Normalization:** The matrix  $\mathbf{Q}_m(w)$  satisfies the condition

$$\lim_{\substack{w \rightarrow \infty \\ |\arg(-w)| < \pi}} \mathbf{Q}_m(w) = \mathbb{I}. \quad (6.8)$$

All of the above jump matrix estimates are uniform for bounded  $m$ .

Note that since the jump matrix for  $\mathbf{Q}_m(w)$  is not specified here precisely, the conditions of Riemann-Hilbert Problem 6.1 are not sufficient to determine  $\mathbf{Q}_m(w)$ , but we shall see that they are enough to approximate  $\mathbf{Q}_m(w)$  with suitable accuracy as  $\epsilon_N \rightarrow 0$ .

Now  $W(\xi)$  is bounded away from zero for  $\xi \in \partial U$ , so from the normalization condition on  $\mathbf{Z}_m(\zeta; y)$  given as part of Riemann-Hilbert Problem 4.1 and understood in the stronger sense of (5.1), and from the fact that  $\eta_m(\xi)$  and its reciprocal are uniformly bounded as  $\epsilon_N \rightarrow 0$  for  $\xi \in \partial U$ , we obtain

$$\dot{\mathbf{P}}_m^{\text{in}}(\xi) \dot{\mathbf{P}}_m^{\text{out}}(\xi)^{-1} = \epsilon_N^{(2m-1)\sigma_3/6} e^{\frac{1}{2}s\sigma_3/\epsilon_N} \left[ \mathbb{I} + \mathcal{O}(\epsilon_N^{1/3}) \right] e^{-\frac{1}{2}s\sigma_3/\epsilon_N} \epsilon_N^{(1-2m)\sigma_3/6} d_N(\xi)^{-\sigma_3}, \quad \xi \in \partial U. \quad (6.9)$$

Recall that  $d_N(w) = 1 + \mathcal{O}(\epsilon_N)$  uniformly for  $w \in \overline{U}$ . It follows that the jump matrix for  $\mathbf{Q}_m(w)$  will be a small perturbation (of order  $\mathcal{O}(\epsilon_N^\mu)$  for some  $\mu \in (0, 1/3)$ ) of the identity for  $\xi \in \partial U$  if

$$\left| \frac{2m-1}{3} - \frac{s}{\epsilon_N \log(\epsilon_N^{-1})} \right| \leq \frac{1}{3} - \mu. \quad (6.10)$$

Subject to this condition, we may formulate a Riemann-Hilbert problem for  $\mathbf{Q}_m(w)$  of small-norm type, that is, a problem that can be solved by iteration. This calculation leads to the suggestion that different choices of the integer  $m$  will be required to control the error in different regions of the  $(y, s)$  parameter plane, an idea that has proven useful in at least two very recent works on related problems of semiclassical analysis for nonlinear waves. In the work of Claeys and Grava [11] on the small-dispersion limit of the KdV equation near the “trailing edge” of the oscillatory wavepacket generated by a gradient catastrophe in the solution of the limiting Burgers equation, the authors find that different choices of an integer parametrizing an outer model problem are required as the rescaled distance from the trailing edge varies. Similarly, in the work of Bertola and Tovbis [4] the semiclassical solution of the focusing NLS equation is examined for points near the “primary caustic” or “breaking curve” generated from an elliptic umbilic catastrophe in the solution of the formal limiting Whitham system, and it is again found that different choices of parametrices must be made as a certain rescaled distance from the caustic curve varies. In both of these problems, as will be the case here, the rescaling of distance involves logarithms of the small parameter to get the balance right, and the resulting approximate solutions have a “solitonic” nature (in the KdV problem they are simple solitons while in the NLS problem they are soliton superpositions forming breathers).

To implement this suggestion, we will choose a value for  $m$  and then define corresponding regions of the  $(y, s)$  plane in which we can make use of this value of  $m$ . Also, we will find it necessary to relax the small-norm condition (6.10) in order to cover a large enough part of the  $(y, s)$  plane with overlapping regions, and this means that the Riemann-Hilbert problem solved by the matrix  $\mathbf{Q}_m(w)$  will no longer be treatable by iteration, and instead we shall have to improve our inadequate model  $\mathbf{P}_m(w)$  by building parametrices for  $\mathbf{Q}_m(w)$ , another technique used in the papers [11] and [4] (also in the more recent work [5] of Bertola and Tovbis where the solution is analyzed in a neighborhood of the elliptic umbilic catastrophe point itself).

For each integer  $m \in \mathbb{Z}$  we define a shifted and scaled coordinate  $p_m$  equivalent to  $s$  by the relation

$$s = \frac{2m}{3} \epsilon_N \log(\epsilon_N^{-1}) + \epsilon_N p_m. \quad (6.11)$$

**Definition 6.1** (Sets  $\Omega_m^\pm$ ). *Let  $m \in \mathbb{Z}$ . For suitably small  $\delta > 0$  and arbitrary  $\kappa \geq 0$ , set*

$$\begin{aligned} \Omega_m^+ := \left\{ (y, s) \in \mathbb{R}^2 \text{ with } |p_m| \leq \frac{1}{3} \log(\epsilon_N^{-1}) + \kappa \text{ such that} \right. \\ \left. \begin{aligned} |y - \mathcal{P}(\mathcal{U}_m)| &\geq \delta e^{\frac{1}{2}p_m} \quad \text{if} \quad -\frac{1}{3} \log(\epsilon_N^{-1}) - \kappa \leq p_m \leq 0 \quad \text{and} \\ |y - \mathcal{Z}(\mathcal{U}_m)| &\leq \delta e^{-\frac{1}{2}p_m} \quad \text{if} \quad 0 < p_m \leq \frac{1}{3} \log(\epsilon_N^{-1}) + \kappa \end{aligned} \right\} \end{aligned} \quad (6.12)$$

and

$$\begin{aligned} \Omega_m^- := \left\{ (y, s) \in \mathbb{R}^2 \text{ with } |p_{m-1}| \leq \frac{1}{3} \log(\epsilon_N^{-1}) + \kappa \text{ such that} \right. \\ \left. \begin{aligned} |y - \mathcal{P}(\mathcal{V}_m)| &\geq \delta e^{-\frac{1}{2}p_{m-1}} \quad \text{if} \quad 0 \leq p_{m-1} \leq \frac{1}{3} \log(\epsilon_N^{-1}) + \kappa \quad \text{and} \\ |y - \mathcal{Z}(\mathcal{V}_m)| &\leq \delta e^{\frac{1}{2}p_{m-1}} \quad \text{if} \quad -\frac{1}{3} \log(\epsilon_N^{-1}) - \kappa \leq p_{m-1} < 0 \end{aligned} \right\}. \end{aligned} \quad (6.13)$$

The set  $\Omega_m^+$  is basically a horizontal strip with excluded *notches* located near the poles of  $\mathcal{U}_m(y)$  and upward-extending *teeth* located near the zeros of  $\mathcal{U}_m(y)$ . Similarly,  $\Omega_m^-$  is a horizontal strip with excluded notches located near the poles of  $\mathcal{V}_m(y)$  and downward-extending teeth located near the zeros of  $\mathcal{V}_m(y)$ . Since the poles of  $\mathcal{U}_m$  and  $\mathcal{V}_m$  coincide (although their zeros do not in general), that is,  $\mathcal{P}(\mathcal{U}_m) = \mathcal{P}(\mathcal{V}_m)$ , the notches of  $\Omega_m^+$  align with those of  $\Omega_m^-$ . This is very important, since we will make use of the solution  $\mathbf{Z}_m(\zeta; y)$  of Riemann-Hilbert Problem 4.1 for  $(y, s) \in \Omega_m^+ \cup \Omega_m^-$ , and therefore we must exclude from this union points with  $y$ -values too close to points of  $\mathcal{P}(\mathcal{U}_m) = \mathcal{P}(\mathcal{V}_m)$  at which  $\mathbf{Z}_m(\zeta; y)$  fails to exist. See Figure 6.2.

Since  $\mathcal{V}_{m+1}(y) = 1/\mathcal{U}_m(y)$  according to (5.62), it is clear that  $\mathcal{P}(\mathcal{U}_{m+1}) = \mathcal{Z}(\mathcal{V}_m)$  and  $\mathcal{Z}(\mathcal{U}_{m+1}) = \mathcal{P}(\mathcal{V}_m)$ . It follows that while the sets  $\Omega_m^+$  and  $\Omega_{m+1}^-$  have disjoint interiors, their union is a horizontal strip:

$$\Omega_m^+ \cup \Omega_{m+1}^- = \left\{ (y, s) \in \mathbb{R}^2 \text{ such that } |p_m| \leq \frac{1}{3} \log(\epsilon_N^{-1}) + \kappa \right\}. \quad (6.14)$$

The two regions interlock near the singularities of  $\log |\mathcal{U}_m(y)|$  (equivalently the singularities of  $\log |\mathcal{V}_{m+1}(y)|$ ); the teeth of  $\Omega_m^+$  fit in a one-to-one fashion into the notches of  $\Omega_{m+1}^-$ , and the teeth of  $\Omega_{m+1}^-$  fit in a one-to-one fashion into the notches of  $\Omega_m^+$ . See Figure 6.3. In Figure 6.4 we illustrate the tiling of a region of the  $(y, s)$ -plane with many such strips when  $\kappa = 0$  (the strips overlap if  $\kappa > 0$ ). This latter figure illustrates another feature of the  $\kappa = 0$  tiling: since by the Bäcklund transformations (5.62)–(5.63),  $y_0$  is a simple zero of  $\mathcal{U}_{m-1}(y)$  if and only if  $y_0$  is a simple pole of both  $\mathcal{U}_m(y)$  and  $\mathcal{V}_m(y)$  if and only if  $y_0$  is a simple zero of  $\mathcal{V}_{m+1}(y)$ , the tips of the teeth of  $\Omega_{m-1}^+$  are matched in a one-to-one fashion onto the tips of the teeth of  $\Omega_{m+1}^-$ .

We will use Definition 6.1 in the following context: we fix a bound  $B > 0$  and consider only indices  $m \in \mathbb{Z}$  with  $|m| \leq B$ ; we then choose  $\delta > 0$  sufficiently small that for all  $m \in \mathbb{Z}$  with  $|m| \leq B$ , no two real

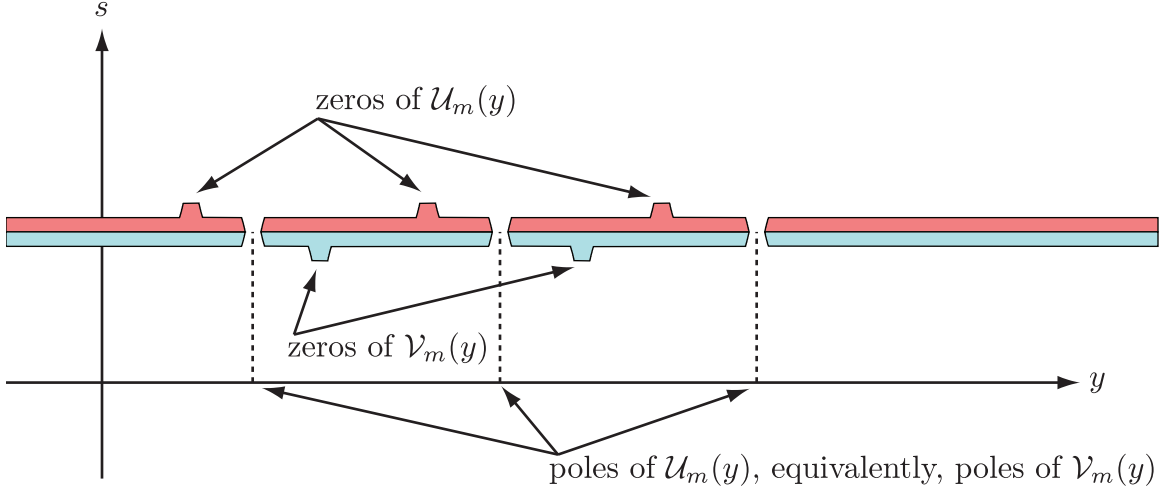


FIGURE 6.2. The regions  $\Omega_m^+$  (pink) and  $\Omega_m^-$  (blue) for  $m = 6$ . For this figure,  $\kappa = 0$ ,  $\delta = 0.2$  and  $\epsilon_N = 0.1$ . Note the alignment of the notches. It is for  $(y, s)$  in the union of these regions that we select the value  $m = 6$  in both the solution  $\mathbf{P}_m^{\text{out}}(w)$  of the outer model problem and the solution  $\mathbf{Z}_m(\zeta; y)$  of the inner model problem.

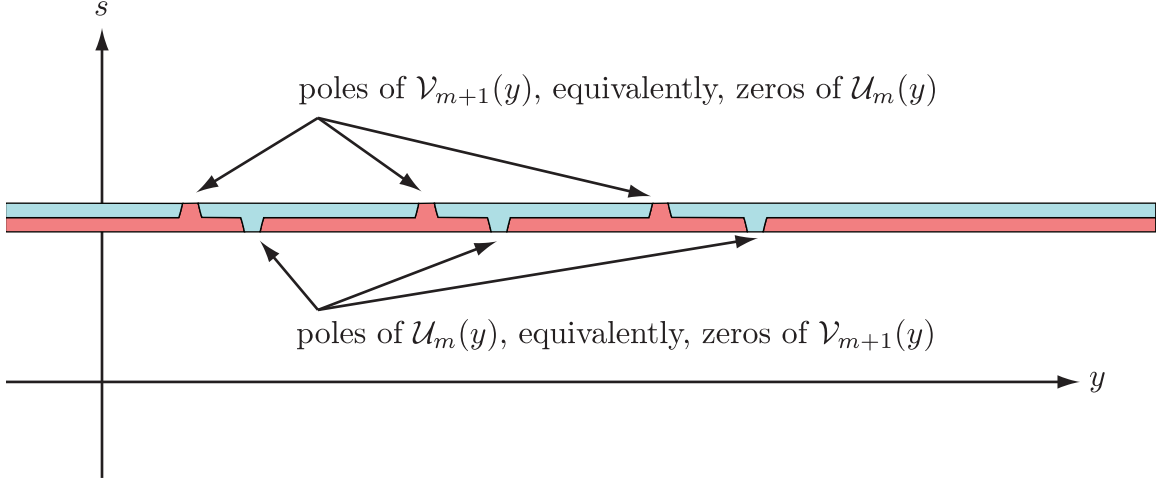


FIGURE 6.3. The regions  $\Omega_m^+$  (pink) and  $\Omega_{m+1}^-$  (blue) for  $m = 6$ . Their union is a closed horizontal strip in the  $(y, s)$ -plane. For this figure,  $\kappa = 0$ ,  $\delta = 0.2$  and  $\epsilon_N = 0.1$ .

singularities of  $\log |\mathcal{U}_m(y)|$  or  $\log |\mathcal{V}_m(y)|$  are closer<sup>1</sup> than  $2\delta$ . This prevents neighboring teeth/notches from overlapping at their widest point (corresponding to  $p_m = 0$ ).

**6.1. Parametrix for  $\mathbf{Q}_m(w)$  in  $\Omega_m^\pm$ .** Even with the choice of the integer  $m$  corresponding to  $(y, s) \in \Omega_m^\pm$  it turns out that Riemann-Hilbert Problem 6.1 cannot generally be solved with the use of small-norm theory as the jump matrix  $\mathbf{V}_{\mathbf{Q}_m}(\xi)$  is not close to the identity for  $\xi \in \partial U$ , at least not uniformly for  $(y, s) \in \Omega_m^\pm$ .

<sup>1</sup>Preliminary asymptotic analysis suggests that  $\delta = \mathcal{O}(B^{-1/3})$  as  $B \uparrow \infty$ , which is one indication of failure of uniformity of the analysis in this paper when  $t \sim s \gg \epsilon_N \log(\epsilon_N^{-1})$ . Indeed, for  $t > 0$  fixed and  $x \approx x_{\text{crit}}$  fixed one expects a different description of the sine-Gordon dynamics for sufficiently small  $\epsilon_N$  in terms of higher-genus Riemann theta functions.

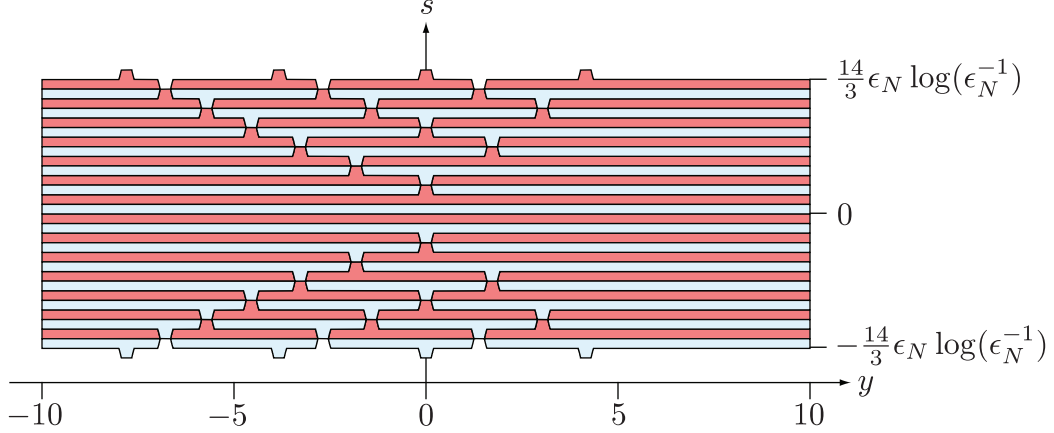


FIGURE 6.4. The regions  $\Omega_{-6}^+, \dots, \Omega_7^+$  (pink) and  $\Omega_{-6}^-, \dots, \Omega_7^-$  (blue), for the “tiling” case  $\kappa = 0$ .

Therefore, even though we may think of  $\mathbf{Q}_m(w)$  as an “error” matrix gauging the mismatch between the matrix function  $\dot{\mathbf{P}}_m(w)$  and the matrix function  $\mathbf{P}(w)$  for which the former was introduced as a model, we must in general carry out further modeling of  $\mathbf{Q}_m(w)$  by creating an explicit model (parametrix) for it, and comparing the model with  $\mathbf{Q}_m(w)$  in hopes of — finally — arriving at a small-norm problem for the matrix ratio. We now proceed to build such a *parametrix for the error* by a procedure referred to in some papers [10, 11] as *improvement of the parametrices*.

Let  $m \in \mathbb{Z}$  with  $|m| \leq B$  be fixed, and suppose that  $(y, s) \in \Omega_m^+$  with  $y \in \mathbb{R}$  bounded. In particular, this implies that the coordinate  $p_m$  defined by (6.11) is subjected to the inequalities  $|p_m| \leq \frac{1}{3} \log(\epsilon_N^{-1}) + \kappa$ . Using the coordinate  $p_m$ , and recalling Proposition 5.2 along with the expansion (5.1), we write the matrix  $\mathbf{V}_{\mathbf{Q}_m}(\xi)$  defined by the right-hand side of (6.4) in the form (assuming only that  $y$  is bounded)

$$\begin{aligned} \mathbf{V}_{\mathbf{Q}_m}(\xi) &= \exp \left( \frac{1}{2} \left( p_m + \frac{1}{3} \log(\epsilon_N^{-1}) \right) \sigma_3 \right) \\ &\quad \cdot \eta_m(\xi)^{\sigma_3} \left[ \mathbb{I} + \epsilon_N^{1/3} \mathbf{A}_m(y) W(\xi)^{-1} + \epsilon_N^{2/3} \mathbf{B}_m(y) W(\xi)^{-2} + \mathcal{O} \left( \frac{\epsilon_N}{|y - \mathcal{P}(\mathcal{U}_m)|} \right) \right] \eta_m(\xi)^{-\sigma_3} \\ &\quad \cdot \exp \left( -\frac{1}{2} \left( p_m + \frac{1}{3} \log(\epsilon_N^{-1}) \right) \sigma_3 \right) d_N(\xi)^{-\sigma_3}, \quad \xi \in \partial U \end{aligned} \quad (6.15)$$

because  $W(\xi)$  is bounded away from zero on  $\partial U$ . Since  $\eta_m(\xi)$ ,  $d_N(\xi)$ , and their reciprocals are all bounded on  $\partial U$  and  $p_m \geq -\frac{1}{3} \log(\epsilon_N^{-1}) - \kappa$  for  $(y, s) \in \Omega_m^+$ ,

$$\begin{aligned} \mathbf{V}_{\mathbf{Q}_m}(\xi) &= \exp \left( \frac{1}{2} \left( p_m + \frac{1}{3} \log(\epsilon_N^{-1}) \right) \sigma_3 \right) \\ &\quad \cdot \eta_m(\xi)^{\sigma_3} \left[ \mathbb{I} + \epsilon_N^{1/3} \mathbf{A}_m(y) W(\xi)^{-1} + \epsilon_N^{2/3} \mathbf{B}_m(y) W(\xi)^{-2} \right] \eta_m(\xi)^{-\sigma_3} \\ &\quad \cdot \exp \left( -\frac{1}{2} \left( p_m + \frac{1}{3} \log(\epsilon_N^{-1}) \right) \sigma_3 \right) d_N(\xi)^{-\sigma_3} + \mathcal{O} \left( \frac{\epsilon_N^{2/3} e^{p_m}}{|y - \mathcal{P}(\mathcal{U}_m)|} \right) \end{aligned} \quad (6.16)$$

holds for  $(y, s) \in \Omega_m^+$  with  $y$  bounded and  $\xi \in \partial U$ . Now, we want to neglect all but the (1, 2)-entry of the matrix  $\epsilon_N^{1/3} \mathbf{A}_m(y) W(\xi)^{-1} + \epsilon_N^{2/3} \mathbf{B}_m(y) W(\xi)^{-2}$ . Since the matrix elements of  $\mathbf{A}_m(y)$  and  $\mathbf{B}_m(y)$  are  $\mathcal{O}(|y - \mathcal{P}(\mathcal{U}_m)|^{-1})$ , we may do this at the cost of adding a new error term of the form  $\mathcal{O}(\epsilon_N^{1/3} |y - \mathcal{P}(\mathcal{U}_m)|^{-1})$  (dominated by the diagonal entries). This error term dominates the one already present in (6.16) for

$(y, s) \in \Omega_m^+$  as a consequence of the inequality  $p_m \leq \frac{1}{3} \log(\epsilon_N^{-1}) + \kappa$ . Therefore,

$$\mathbf{V}_{\mathbf{Q}_m}(\xi) = \begin{bmatrix} 1 & Y_m^+(\xi) \\ 0 & 1 \end{bmatrix} d_N(\xi)^{-\sigma_3} + \mathcal{O}\left(\frac{\epsilon_N^{1/3}}{|y - \mathcal{P}(\mathcal{U}_m)|}\right), \quad \xi \in \partial U, \quad (y, s) \in \Omega_m^+ \text{ with } y \text{ bounded}, \quad (6.17)$$

where, defining

$$h_m^+(w) := \eta_m(w)^2, \quad (6.18)$$

the explicit function appearing in (6.17) is

$$\begin{aligned} Y_m^+(w) &:= e^{p_m} h_m^+(w) \left( A_{m,12}(y) W(w)^{-1} + \epsilon_N^{1/3} B_{m,12}(y) W(w)^{-2} \right) \\ &= e^{p_m} h_m^+(w) \left( \mathcal{U}_m(y) W(w)^{-1} + \epsilon_N^{1/3} [2H_m(y) \mathcal{U}_m(y) - \mathcal{U}_m'(y)] W(w)^{-2} \right). \end{aligned} \quad (6.19)$$

Note that  $Y_m^+(w)$  extends from  $\xi \in \partial U$  as a meromorphic function of  $w \in \bar{U}$  whose only singularity is a double pole at  $w = w_*$ , and that  $Y_m^+(\xi)$  also depends parametrically on  $p_m$  and  $y$ . Since  $d_N(\xi) = 1 + \mathcal{O}(\epsilon_N)$  holds uniformly for  $\xi \in \partial U$ , another way to write (6.17) is

$$\begin{aligned} \mathbf{V}_{\mathbf{Q}_m}(\xi) \begin{bmatrix} 1 & Y_m^+(\xi) \\ 0 & 1 \end{bmatrix}^{-1} &= \mathbb{I} + \mathcal{O}\left(\frac{\epsilon_N^{1/3}}{|y - \mathcal{P}(\mathcal{U}_m)|}\right) + \mathcal{O}\left(\frac{Y_m^+(\xi) \epsilon_N^{1/3}}{|y - \mathcal{P}(\mathcal{U}_m)|}\right), \\ &\quad \xi \in \partial U, \quad (y, s) \in \Omega_m^+ \text{ with } y \text{ bounded}. \end{aligned} \quad (6.20)$$

**Proposition 6.1.** *The following three estimates hold uniformly for  $\xi \in \partial U$  and  $y$  bounded.*

*In the upward-pointing “teeth” of  $\Omega_m^+$ :*

$$\begin{aligned} \mathbf{V}_{\mathbf{Q}_m}(\xi) \begin{bmatrix} 1 & Y_m^+(\xi) \\ 0 & 1 \end{bmatrix}^{-1} &= \mathbb{I} + \mathcal{O}(\epsilon_N^{1/3} e^{p_m} |y - \mathcal{Z}(\mathcal{U}_m)|), \\ &\quad (y, s) \in \Omega_m^+, \quad p_m > 0 \text{ and } |y - \mathcal{Z}(\mathcal{U}_m)| \geq \delta e^{-p_m}. \end{aligned} \quad (6.21)$$

*In the part of  $\Omega_m^+$  resembling a “strip with notches”:*

$$\mathbf{V}_{\mathbf{Q}_m}(\xi) \begin{bmatrix} 1 & Y_m^+(\xi) \\ 0 & 1 \end{bmatrix}^{-1} = \mathbb{I} + \mathcal{O}\left(\frac{\epsilon_N^{1/3}}{|y - \mathcal{P}(\mathcal{U}_m)|}\right), \quad (y, s) \in \Omega_m^+, \quad |y - \mathcal{P}(\mathcal{U}_m)| \leq \delta. \quad (6.22)$$

*Elsewhere in  $\Omega_m^+$ :*

$$\mathbf{V}_{\mathbf{Q}_m}(\xi) \begin{bmatrix} 1 & Y_m^+(\xi) \\ 0 & 1 \end{bmatrix}^{-1} = \mathbb{I} + \mathcal{O}(\epsilon_N^{1/3}). \quad (6.23)$$

*Proof.* First suppose that  $0 < p_m \leq \frac{1}{3} \log(\epsilon_N^{-1}) + \kappa$  (so we are considering  $(y, s)$  to lie in the upward-pointing “teeth” of  $\Omega_m^+$ ) and that  $y$  is bounded. In this sub-region,  $|y - \mathcal{P}(\mathcal{U}_m)|$  is bounded away from zero because the teeth are localized near the zeros  $\mathcal{Z}(\mathcal{U}_m)$  of  $\mathcal{U}_m(y)$ . Therefore  $2H_m \mathcal{U}_m - \mathcal{U}_m'$  is bounded and  $\mathcal{U}_m(y) = \mathcal{O}(|y - \mathcal{Z}(\mathcal{U}_m)|)$ , so  $Y_m^+(\xi) = \mathcal{O}(e^{p_m} |y - \mathcal{Z}(\mathcal{U}_m)|) + \mathcal{O}(1)$  holds uniformly for  $\xi \in \partial U$ . Putting this estimate into (6.20) gives

$$\begin{aligned} \mathbf{V}_{\mathbf{Q}_m}(\xi) \begin{bmatrix} 1 & Y_m^+(\xi) \\ 0 & 1 \end{bmatrix}^{-1} &= \mathbb{I} + \mathcal{O}(\epsilon_N^{1/3}) + \mathcal{O}(\epsilon_N^{1/3} e^{p_m} |y - \mathcal{Z}(\mathcal{U}_m)|), \\ &\quad \xi \in \partial U, \quad (y, s) \in \Omega_m^+ \text{ with } y \text{ bounded and } p_m > 0. \end{aligned} \quad (6.24)$$

This proves the estimate (6.21), and it also proves the estimate (6.23) in the case that  $p_m > 0$ .

Now suppose instead that  $-\frac{1}{3} \log(\epsilon_N^{-1}) - \kappa \leq p_m \leq 0$  (so we are considering  $(y, s)$  to lie in the complementary part of  $\Omega_m^+$  that resembles a “strip with notches”), and that  $y$  is bounded. Since according to Proposition 5.2, both  $\mathcal{U}_m(y)$  and  $2H_m(y) \mathcal{U}_m(y) - \mathcal{U}_m'(y)$  have simple poles at the points of  $\mathcal{P}(\mathcal{U}_m)$ , it follows that in this sub-region,  $Y_m^+(\xi) = \mathcal{O}(e^{p_m} |y - \mathcal{P}(\mathcal{U}_m)|^{-1})$ . But in this subregion we also have the inequality

$|y - \mathcal{P}(\mathcal{U}_m)| \geq \delta e^{\frac{1}{2}p_m}$ , so with  $p_m \leq 0$  in fact we have  $Y_m^+(\xi) = \mathcal{O}(1)$ . Putting this information into (6.20) gives

$$\mathbf{V}_{\mathbf{Q}}(\xi) \begin{bmatrix} 1 & Y_m^+(\xi) \\ 0 & 1 \end{bmatrix}^{-1} = \mathbb{I} + \mathcal{O}\left(\frac{\epsilon_N^{1/3}}{|y - \mathcal{P}(\mathcal{U}_m)|}\right), \quad \xi \in \partial U, \quad (y, s) \in \Omega_m^+ \text{ with } y \text{ bounded and } p_m \leq 0. \quad (6.25)$$

This proves (6.22), and it also proves (6.23) in the case that  $p_m \leq 0$ .  $\square$

**Corollary 6.1.** *The following estimate holds uniformly for  $\xi \in \partial U$  and  $(y, s) \in \Omega_m^+$  with  $y$  bounded:*

$$\mathbf{V}_{\mathbf{Q}_m}(\xi) \begin{bmatrix} 1 & Y_m^+(\xi) \\ 0 & 1 \end{bmatrix}^{-1} = \mathbb{I} + \mathcal{O}(\epsilon_N^{1/6}). \quad (6.26)$$

While uniform, this estimate fails to be sharp except possibly at the vertical extremes of  $\Omega_m^+$  where  $|p_m| \sim \frac{1}{3} \log(\epsilon_N^{-1})$ .

As  $\epsilon_N \downarrow 0$ , the function  $Y_m^+(\xi)$  is not uniformly small on  $\partial U$  as  $(y, s)$  ranges over  $\Omega_m^+$  with  $y$  bounded. This implies that to approximate  $\mathbf{Q}_m(w)$  for  $(y, s) \in \Omega_m^+$ , it will be necessary to build a parametrix for it that takes into account the explicit terms in (6.17).

**Riemann-Hilbert Problem 6.2** (Parametrix for  $\mathbf{Q}_m(w)$  in  $\Omega_m^+$ ). *Let real numbers  $y$  and  $p_m$  be given, and assume that  $y \notin \mathcal{P}(\mathcal{U}_m)$ . Seek a matrix  $\dot{\mathbf{Q}}_m^+(w)$  with the following properties:*

**Analyticity:**  $\dot{\mathbf{Q}}_m^+(w)$  is analytic for  $w \in \mathbb{C} \setminus (\mathbb{R}_+ \cup \partial U)$ , and is uniformly Hölder continuous up to the boundary of its domain of analyticity.

**Jump conditions:** The boundary values taken by  $\dot{\mathbf{Q}}_m^+(w)$  on the contour  $\mathbb{R}_+ \cup \partial U$  are related as follows:

$$\dot{\mathbf{Q}}_{m+}^+(\xi) = \sigma_2 \dot{\mathbf{Q}}_{m-}^+(\xi) \sigma_2, \quad \xi > 0, \quad (6.27)$$

and

$$\dot{\mathbf{Q}}_{m+}^+(\xi) = \dot{\mathbf{Q}}_{m-}^+(\xi) \begin{bmatrix} 1 & Y_m^+(\xi) \\ 0 & 1 \end{bmatrix}, \quad \xi \in \partial U \quad (6.28)$$

where  $\partial U$  is taken to be negatively (clockwise) oriented and  $Y_m^+(\xi)$  is defined by (6.19).

**Normalization:** the matrix  $\dot{\mathbf{Q}}_m^+(w)$  satisfies the condition

$$\lim_{w \rightarrow \infty} \dot{\mathbf{Q}}_m^+(w) = \mathbb{I}. \quad (6.29)$$

We solve this problem uniquely for  $\dot{\mathbf{Q}}_m^+(w)$  as follows. Firstly, define a related unknown  $\mathbf{F}_m^+(w)$  as follows:

$$\mathbf{F}_m^+(w) := \begin{cases} \dot{\mathbf{Q}}_m^+(w), & w \notin \overline{U} \\ \dot{\mathbf{Q}}_m^+(w) \begin{bmatrix} 1 & Y_m^+(w) \\ 0 & 1 \end{bmatrix}, & w \in U. \end{cases} \quad (6.30)$$

Clearly,  $\mathbf{F}_m^+(w)$  is continuous across the contour  $\partial U$ . Now,  $\eta_m(w)$  is analytic in  $U$ , but  $W(w)$  has a simple zero at  $w = w_* \approx -1$ , so  $\mathbf{F}_m^+(w)$  generally has a double pole at  $w = w_*$ . To handle the only remaining jump condition for  $\mathbf{F}_m^+(w)$ , namely  $\mathbf{F}_{m+}^+(\xi) = \sigma_2 \mathbf{F}_{m-}^+(\xi) \sigma_2$  for  $\xi > 0$ , we introduce  $\tau := i(-w)^{1/2}$  and set

$$\mathbf{G}_m^+(\tau) := \begin{cases} \mathbf{F}_m^+(\tau^2), & \Im\{\tau\} > 0 \\ \sigma_2 \mathbf{F}_m^+(\tau^2) \sigma_2, & \Im\{\tau\} < 0. \end{cases} \quad (6.31)$$

It is now evident that  $\mathbf{G}_m^+(\tau)$  is a rational function of  $\tau \in \mathbb{C}$  with simple poles at  $\tau = \pm iS$  where  $S := (-w_*)^{1/2} > 0$ , normalized to the identity at  $\tau = \infty$ , and satisfying the symmetry relation  $\mathbf{G}_m^+(-\tau) = \sigma_2 \mathbf{G}_m^+(\tau) \sigma_2$ .

To fully characterize the singularities of  $\mathbf{G}_m^+(\tau)$ , we will relate its principal part at  $\tau = iS$  to its regular part. Note that in a neighborhood of  $\tau = iS$ ,  $Y_m^+(\tau^2)$  has the Laurent expansion

$$Y_m^+(\tau^2) = \frac{J_m^+}{(\tau - iS)^2} + \frac{iK_m^+}{\tau - iS} + \mathcal{O}(1) \quad (6.32)$$

where  $\mathcal{O}(1)$  represents a function analytic at  $\tau = iS$  and where

$$\begin{aligned} J_m^+ &:= \frac{e^{p_m} h_m^+(w_*) \epsilon_N^{1/3}}{4w_* W'(w_*)^2} \left( 2H_m(y) \mathcal{U}_m(y) - \mathcal{U}_m'(y) \right) \\ K_m^+ &:= -\frac{e^{p_m}}{4w_* S W'(w_*)^3} \left[ 2w_* h_m^+(w_*) W'(w_*)^2 \mathcal{U}_m(y) \right. \\ &\quad \left. + \epsilon_N^{1/3} \left( 2H_m(y) \mathcal{U}_m(y) - \mathcal{U}_m'(y) \right) \left( 2w_* h_m^{+'}(w_*) W'(w_*) - 2w_* h_m^+(w_*) W''(w_*) - h_m^+(w_*) W'(w_*) \right) \right]. \end{aligned} \quad (6.33)$$

Note also that  $J_m^+$  and  $K_m^+$  are both real when  $y \in \mathbb{R}$  and  $p_m \in \mathbb{R}$ . Let us write the Laurent expansion of  $\mathbf{G}_m^+(\tau)$  about  $\tau = iS$  in the form

$$\mathbf{G}_m^+(\tau) = \frac{(\mathbf{a}_m^+, \mathbf{b}_m^+)}{(\tau - iS)^2} + \frac{(\mathbf{c}_m^+, \mathbf{d}_m^+)}{\tau - iS} + (\mathbf{e}_m^+, \mathbf{f}_m^+) + (\mathbf{g}_m^+, \mathbf{h}_m^+)(\tau - iS) + \mathcal{O}((\tau - iS)^2), \quad (6.34)$$

where  $\mathcal{O}((\tau - iS)^2)$  represents a matrix function analytic and vanishing to second order at  $\tau = iS$ . Since  $\dot{\mathbf{Q}}_m^+(w)$  is meant to be analytic at  $w = w_*$ , from (6.30) we see that the principal part of  $\mathbf{G}_m^+(\tau)$  at  $\tau = iS$  is determined from its holomorphic part as follows. The first column vanishes:

$$\mathbf{a}_m^+ = \mathbf{c}_m^+ = \mathbf{0}, \quad (6.35)$$

and the second column satisfies the linear relations

$$\mathbf{b}_m^+ = J_m^+ \mathbf{e}_m^+ \quad \text{and} \quad \mathbf{d}_m^+ = J_m^+ \mathbf{g}_m^+ + iK_m^+ \mathbf{e}_m^+. \quad (6.36)$$

Taking into account the symmetry condition  $\mathbf{G}_m^+(-\tau) = \sigma_2 \mathbf{G}_m^+(\tau) \sigma_2$  and the notation of the principal part of the Laurent expansion (6.34), the matrix  $\mathbf{G}_m^+(\tau)$  necessarily can be expressed in partial fractions as follows:

$$\mathbf{G}_m^+(\tau) = \mathbb{I} + \frac{(\mathbf{0}, \mathbf{b}_m^+)}{(\tau - iS)^2} + \frac{(\mathbf{0}, \mathbf{d}_m^+)}{\tau - iS} + \frac{(i\sigma_2 \mathbf{b}_m^+, \mathbf{0})}{(\tau + iS)^2} + \frac{(-i\sigma_2 \mathbf{d}_m^+, \mathbf{0})}{\tau + iS}, \quad (6.37)$$

where we have also used (6.35). Evaluating the first column of this expression and its derivative at the regular point (for the first column)  $\tau = iS$  we find that, in the notation of the regular part of the Laurent expansion (6.34),

$$\mathbf{e}_m^+ = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \frac{i\sigma_2}{4w_*} \mathbf{b}_m^+ - \frac{\sigma_2}{2S} \mathbf{d}_m^+ \quad \text{and} \quad \mathbf{g}_m^+ = -\frac{\sigma_2}{4w_* S} \mathbf{b}_m^+ + \frac{i\sigma_2}{4w_*} \mathbf{d}_m^+, \quad S := (-w_*)^{1/2} > 0. \quad (6.38)$$

Substituting these expressions into the relations (6.36) yields a closed inhomogeneous linear system for the vectors  $\mathbf{b}_m^+$  and  $\mathbf{d}_m^+$ :

$$\begin{aligned} \left( \mathbb{I} - \frac{iJ_m^+}{4w_*} \sigma_2 \right) \mathbf{b}_m^+ + \frac{J_m^+}{2S} \sigma_2 \mathbf{d}_m^+ &= \begin{bmatrix} J_m^+ \\ 0 \end{bmatrix} \\ \frac{J_m^+ + SK_m^+}{4w_* S} \sigma_2 \mathbf{b}_m^+ + \left( \mathbb{I} - \frac{iJ_m^+ + 2iSK_m^+}{4w_*} \sigma_2 \right) \mathbf{d}_m^+ &= \begin{bmatrix} iK_m^+ \\ 0 \end{bmatrix}. \end{aligned} \quad (6.39)$$

This system has a unique solution:

$$\begin{aligned} \mathbf{b}_m^+ &= \frac{4w_* J_m^+}{D_m^+} \begin{bmatrix} 4w_*(16w_*^2 + 2SJ_m^+ K_m^+ + 3[J_m^+]^2) \\ [J_m^+]^3 - 16w_*^2 J_m^+ - 32w_*^2 SK_m^+ \end{bmatrix} \\ \mathbf{d}_m^+ &= \frac{4iS}{D_m^+} \begin{bmatrix} 4w_*(2[J_m^+]^3 + S[J_m^+]^2 K_m^+ - 16w_*^2 SK_m^+) \\ [J_m^+]^4 + 16w_*^2 [J_m^+]^2 + 32w_*^2 SJ_m^+ K_m^+ - 32w_*^3 [K_m^+]^2 \end{bmatrix}, \end{aligned} \quad (6.40)$$

where the denominator

$$D_m^+ := (16w_*^2 + [J_m^+]^2)^2 + 64w_*^2 (J_m^+ + SK_m^+)^2 \geq 256w_*^4 \quad (6.41)$$

is strictly positive, because  $w_* \approx -1$ . Substituting (6.40)–(6.41) into the partial fraction representation (6.37) completes the construction of  $\mathbf{G}_m^+(\tau)$ . By writing  $\tau = i(-w)^{1/2}$  and restricting  $\mathbf{G}_m^+(\tau)$  to the upper half-plane we then recover  $\mathbf{F}_m^+(w)$ , and then from (6.30) we recover the solution  $\dot{\mathbf{Q}}_m^+(w)$  of Riemann-Hilbert Problem 6.2.

**Proposition 6.2.** *For each fixed integer  $m$  with  $|m| \leq B$ , Riemann-Hilbert Problem 6.2 has a unique and explicit solution  $\dot{\mathbf{Q}}_m^+(w)$  satisfying  $\det(\dot{\mathbf{Q}}_m^+(w)) = 1$  where defined. Also,  $\dot{\mathbf{Q}}_m^+(w)$  is uniformly bounded for  $w \in \mathbb{C} \setminus (U \cup \mathbb{R}_+)$  and  $(y, s) \in \Omega_m^+$  with  $y$  bounded as  $\epsilon_N \rightarrow 0$ .*

*Proof.* The above construction clearly yields a solution of Riemann-Hilbert Problem 6.2, and uniqueness follows easily from the conditions of that problem via a Liouville argument, which also shows that  $\dot{\mathbf{Q}}_m^+(w)$  is unimodular. Since  $\dot{\mathbf{Q}}_m^+(w) = \mathbf{F}_m^+(w)$  for  $w \in \mathbb{C} \setminus U$ , the uniform bound on  $\dot{\mathbf{Q}}_m^+(w)$  will follow from a corresponding bound on  $\mathbf{F}_m^+(w)$  valid when  $w$  is bounded away from the singularity at  $w_*$ , or what is the same, a bound on  $\mathbf{G}_m^+(\tau)$  valid when  $\tau$  is bounded away from both  $iS$  and  $-iS$  where  $S := (-w_*)^{1/2} > 0$ . According to the partial-fraction representation (6.37) of  $\mathbf{G}_m^+(\tau)$ , it is therefore enough to establish that the elements of the vectors  $\mathbf{b}_m^+$  and  $\mathbf{d}_m^+$  remain bounded as  $\epsilon_N \downarrow 0$  uniformly with respect to  $(y, s) \in \Omega_m^+$  with  $y$  bounded. But from (6.40) we see that the components of  $\mathbf{b}_m^+$  and  $\mathbf{d}_m^+$  really only depend on  $J_m^+$  and  $K_m^+$  (since  $w_* \approx -1$ ) and while the latter can become unbounded, it is still obvious that  $\mathbf{b}_m^+$  and  $\mathbf{d}_m^+$  cannot.  $\square$

**6.2. Parametrix for  $\mathbf{Q}_m(w)$  in  $\Omega_m^-$ .** Fix an integer  $m$  with  $|m| \leq B$ , and suppose that  $(y, s) \in \Omega_m^-$  with  $y \in \mathbb{R}$  bounded. By introducing the scaled and shifted coordinate  $p_{m-1}$  defined by (6.11) and analyzing the jump matrix  $\mathbf{V}_{\mathbf{Q}_m}(\xi)$  for  $\xi \in \partial U$  as in the beginning of §6.1, we see that in the current situation a lower-triangular model for the jump matrix is more appropriate. In particular, defining

$$h_m^-(w) := \eta_m(w)^{-2} \quad (6.42)$$

and then

$$Y_m^-(w) := e^{-p_{m-1}} h_m^-(w) \left( \mathcal{V}_m(y) W(w)^{-1} + \epsilon_N^{1/3} [\mathcal{V}'_m(y) - 2H_m(y) \mathcal{V}_m(y)] W(w)^{-2} \right), \quad (6.43)$$

(note that  $Y_m^-(\xi)$  extends from  $\xi \in \partial U$  to a meromorphic function for  $\xi \in \bar{U}$  whose only singularity is a double pole at  $w = w_*$ ) we arrive at the following.

**Proposition 6.3.** *The following three estimates hold uniformly for  $\xi \in \partial U$  and  $y$  bounded. In the downward-pointing “teeth” of  $\Omega_m^-$ :*

$$\mathbf{V}_{\mathbf{Q}_m}(\xi) \begin{bmatrix} 1 & 0 \\ Y_m^-(\xi) & 1 \end{bmatrix}^{-1} = \mathbb{I} + \mathcal{O}(\epsilon_N^{1/3} e^{-p_{m-1}} |y - \mathcal{Z}(\mathcal{V}_m)|), \quad (y, s) \in \Omega_m^-, \quad p_{m-1} < 0 \text{ and } |y - \mathcal{Z}(\mathcal{V}_m)| \geq \delta e^{p_{m-1}}. \quad (6.44)$$

*In the part of  $\Omega_m^-$  resembling a “strip with notches”:*

$$\mathbf{V}_{\mathbf{Q}_m}(\xi) \begin{bmatrix} 1 & 0 \\ Y_m^-(\xi) & 1 \end{bmatrix}^{-1} = \mathbb{I} + \mathcal{O}\left(\frac{\epsilon_N^{1/3}}{|y - \mathcal{P}(\mathcal{V}_m)|}\right), \quad (y, s) \in \Omega_m^-, \quad |y - \mathcal{P}(\mathcal{V}_m)| \leq \delta. \quad (6.45)$$

*Elsewhere in  $\Omega_m^-$ :*

$$\mathbf{V}_{\mathbf{Q}_m}(\xi) \begin{bmatrix} 1 & 0 \\ Y_m^-(\xi) & 1 \end{bmatrix}^{-1} = \mathbb{I} + \mathcal{O}(\epsilon_N^{1/3}). \quad (6.46)$$

The proof of this statement is virtually the same as that of Proposition 6.1. The analogue for  $\Omega_m^-$  of Corollary 6.1 is then this result.

**Corollary 6.2.** *The following estimate holds uniformly for  $\xi \in \partial U$  and  $(y, s) \in \Omega_m^-$  with  $y$  bounded:*

$$\mathbf{V}_{\mathbf{Q}_m}(\xi) \begin{bmatrix} 1 & 0 \\ Y_m^-(\xi) & 1 \end{bmatrix}^{-1} = \mathbb{I} + \mathcal{O}(\epsilon_N^{1/6}). \quad (6.47)$$

While uniform, this estimate fails to be sharp except possibly at the vertical extremes of  $\Omega_m^-$  where  $|p_{m-1}| \sim \frac{1}{3} \log(\epsilon_N^{-1})$ .

Based on this approximation result for  $\mathbf{V}_{\mathbf{Q}_m}(\xi)$ , we build a parametrix for  $\mathbf{Q}_m(w)$  to satisfy the following problem.

**Riemann-Hilbert Problem 6.3** (Parametrix for  $\mathbf{Q}_m(w)$  in  $\Omega_m^-$ ). *Let real numbers  $y$  and  $p_{m-1}$  be given, and assume that  $y \notin \mathcal{P}(\mathcal{V}_m)$ . Seek a matrix  $\dot{\mathbf{Q}}_m^-(w)$  with the following properties:*



**Analyticity:**  $\dot{\mathbf{Q}}_m^-(w)$  is analytic for  $w \in \mathbb{C} \setminus (\mathbb{R}_+ \cup \partial U)$ , and is uniformly Hölder continuous up to the boundary of its domain of analyticity.

**Jump conditions:** The boundary values taken by  $\dot{\mathbf{Q}}_m^-(w)$  on the contour  $\mathbb{R}_+ \cup \partial U$  are related as follows:

$$\dot{\mathbf{Q}}_{m+}^-(\xi) = \sigma_2 \dot{\mathbf{Q}}_{m-}^-(\xi) \sigma_2, \quad \xi > 0, \quad (6.48)$$

and

$$\dot{\mathbf{Q}}_{m+}^-(\xi) = \dot{\mathbf{Q}}_{m-}^-(\xi) \begin{bmatrix} 1 & 0 \\ Y_m^-(\xi) & 1 \end{bmatrix}, \quad \xi \in \partial U \quad (6.49)$$

where  $\partial U$  is taken to be negatively (clockwise) oriented and  $Y_m^-(\xi)$  is defined by (6.43).

**Normalization:** the matrix  $\dot{\mathbf{Q}}_m^-(w)$  satisfies the condition

$$\lim_{w \rightarrow \infty} \dot{\mathbf{Q}}_m^-(w) = \mathbb{I}. \quad (6.50)$$

This problem is solved completely analogously to Riemann-Hilbert Problem 6.2. Namely, we introduce a matrix  $\mathbf{G}_m^-(\tau)$  by

$$\mathbf{G}_m^-(\tau) := \begin{cases} \mathbf{F}_m^-(\tau^2), & \Im\{\tau\} > 0 \\ \sigma_2 \mathbf{F}_m^-(\tau^2) \sigma_2, & \Im\{\tau\} < 0, \end{cases} \quad (6.51)$$

where we define  $\mathbf{F}_m^-(w)$  by

$$\mathbf{F}_m^-(w) := \begin{cases} \dot{\mathbf{Q}}_m^-(w), & w \notin \overline{U} \\ \dot{\mathbf{Q}}_m^-(w) \begin{bmatrix} 1 & 0 \\ Y_m^-(w) & 1 \end{bmatrix}, & w \in U. \end{cases} \quad (6.52)$$

Then,  $\mathbf{G}_m^-(\tau)$  is a rational function with  $\mathbf{G}_m^-(\infty) = \mathbb{I}$  and double poles at  $\tau = \pm iS$  where  $S := (-w_*)^{1/2} > 0$  only, satisfying the symmetry relation  $\mathbf{G}_m^-(-\tau) = \sigma_2 \mathbf{G}_m^-(\tau) \sigma_2$ .

The singularities of  $\mathbf{G}_m^-(\tau)$  are characterized by first noting the Laurent expansion of  $Y_m^-(\tau^2)$  about  $\tau = iS$ :

$$Y_m^-(\tau^2) = \frac{J_m^-}{(\tau - iS)^2} + \frac{iK_m^-}{\tau - iS} + \mathcal{O}(1) \quad (6.53)$$

where  $\mathcal{O}(1)$  denotes a function analytic at  $\tau = iS$  and where

$$\begin{aligned} J_m^- &:= -\frac{e^{-p_{m-1}} h_m^-(w_*) \epsilon_N^{1/3}}{4w_* W'(w_*)^2} \left( 2H_m(y) \mathcal{V}_m(y) - \mathcal{V}_m'(y) \right) \\ K_m^- &:= -\frac{e^{-p_{m-1}}}{4w_* S W'(w_*)^3} \left[ 2w_* h_m^-(w_*) W'(w_*)^2 \mathcal{V}_m(y) \right. \\ &\quad \left. + \epsilon_N^{1/3} \left( 2H_m(y) \mathcal{V}_m(y) - \mathcal{V}_m'(y) \right) \left( -2w_* h_m^-(w_*) W'(w_*) + 2w_* h_m^-(w_*) W''(w_*) + h_m^-(w_*) W'(w_*) \right) \right] \end{aligned} \quad (6.54)$$

are real-valued coefficients when both  $y \in \mathbb{R}$  and  $p_{m-1} \in \mathbb{R}$ . The Laurent series of  $\mathbf{G}_m^-(\tau)$  about  $\tau = iS$  has the form

$$\mathbf{G}_m^-(\tau) = \frac{(\mathbf{a}_m^-, \mathbf{b}_m^-)}{(\tau - iS)^2} + \frac{(\mathbf{c}_m^-, \mathbf{d}_m^-)}{\tau - iS} + (\mathbf{e}_m^-, \mathbf{f}_m^-) + (\mathbf{g}_m^-, \mathbf{h}_m^-)(\tau - iS) + \mathcal{O}((\tau - iS)^2), \quad (6.55)$$

where  $\mathcal{O}((\tau - iS)^2)$  represents an analytic function vanishing to second order at  $\tau = iS$ . Since  $\dot{\mathbf{Q}}_m^-(w)$  is meant to be analytic at  $w = w_*$ , from (6.52) we learn that the principal part of the second column of  $\mathbf{G}_m^-(\tau)$  vanishes:

$$\mathbf{b}_m^- = \mathbf{d}_m^- = \mathbf{0}, \quad (6.56)$$

and the principal part of the first column satisfies the relations

$$\mathbf{a}_m^- = J_m^- \mathbf{f}_m^- \quad \text{and} \quad \mathbf{c}_m^- = J_m^- \mathbf{h}_m^- + iK_m^- \mathbf{f}_m^-. \quad (6.57)$$

Using the symmetry condition  $\mathbf{G}_m^-(\tau) = \sigma_2 \mathbf{G}_m^-(\tau) \sigma_2$  and taking into account (6.56), we see that the partial fraction expansion of  $\mathbf{G}_m^-(\tau)$  necessarily has the form:

$$\mathbf{G}_m^-(\tau) = \mathbb{I} + \frac{(\mathbf{a}_m^-, \mathbf{0})}{(\tau - iS)^2} + \frac{(\mathbf{c}_m^-, \mathbf{0})}{\tau - iS} + \frac{(\mathbf{0}, -i\sigma_2 \mathbf{a}_m^-)}{(\tau + iS)^2} + \frac{(\mathbf{0}, i\sigma_2 \mathbf{c}_m^-)}{\tau + iS}. \quad (6.58)$$

By evaluating the second column of this expression and its derivative at  $\tau = iS$  we can express  $\mathbf{f}_m^-$  and  $\mathbf{h}_m^-$  in terms of  $\mathbf{a}_m^-$  and  $\mathbf{c}_m^-$ :

$$\mathbf{f}_m^- = \begin{bmatrix} 0 \\ 1 \end{bmatrix} - \frac{i\sigma_2}{4w_*} \mathbf{a}_m^- + \frac{\sigma_2}{2S} \mathbf{c}_m^- \quad \text{and} \quad \mathbf{h}_m^- = \frac{\sigma_2}{4w_* S} \mathbf{a}_m^- - \frac{i\sigma_2}{4w_*} \mathbf{c}_m^-, \quad S := (-w_*)^{1/2} > 0. \quad (6.59)$$

Combining this with (6.57) yields a closed inhomogeneous linear system of equations for the components of the vectors  $\mathbf{a}_m^-$  and  $\mathbf{c}_m^-$ :

$$\begin{aligned} \left( \mathbb{I} + \frac{iJ_m^-}{4w_*} \sigma_2 \right) \mathbf{a}_m^- - \frac{J_m^-}{2S} \sigma_2 \mathbf{c}_m^- &= \begin{bmatrix} 0 \\ J_m^- \end{bmatrix} \\ -\frac{J_m^- + SK_m^-}{4w_* S} \sigma_2 \mathbf{a}_m^- + \left( \mathbb{I} + \frac{iJ_m^- + 2iSK_m^-}{4w_*} \sigma_2 \right) \mathbf{c}_m^- &= \begin{bmatrix} 0 \\ iK_m^- \end{bmatrix}. \end{aligned} \quad (6.60)$$

As before, this system has a unique solution:

$$\begin{aligned} \mathbf{a}_m^- &= \frac{4w_* J_m^-}{D_m^-} \begin{bmatrix} [J_m^-]^3 - 16w_*^2 J_m^- - 32w_*^2 SK_m^- \\ 4w_*(16w_*^2 + 2SJ_m^- K_m^- + 3[J_m^-]^2) \end{bmatrix} \\ \mathbf{c}_m^- &= \frac{4iS}{D_m^-} \begin{bmatrix} [J_m^-]^4 + 16w_*^2 [J_m^-]^2 + 32w_*^2 SJ_m^- K_m^- - 32w_*^3 [K_m^-]^2 \\ 4w_*(2[J_m^-]^3 + S[J_m^-]^2 K_m^- - 16w_*^2 SK_m^-) \end{bmatrix} \end{aligned} \quad (6.61)$$

where the denominator

$$D_m^- := (16w_*^2 + [J_m^-]^2)^2 + 64w_*^2 (J_m^- + SK_m^-)^2 \geq 256w_*^4 \quad (6.62)$$

is again strictly positive because  $w_* \approx -1$ . This essentially completes the construction of  $\dot{\mathbf{Q}}_m^-(w)$  solving Riemann-Hilbert Problem 6.3.

The analogue of Proposition 6.2 in this case is the following. (Its proof is also nearly the same.)

**Proposition 6.4.** *For each fixed integer  $m$  with  $|m| \leq B$ , Riemann-Hilbert Problem 6.3 has a unique and explicit solution  $\dot{\mathbf{Q}}_m^-(w)$  satisfying  $\det(\dot{\mathbf{Q}}_m^-(w)) = 1$  where defined. Also,  $\dot{\mathbf{Q}}_m^-(w)$  is uniformly bounded for  $w \in \mathbb{C} \setminus (U \cup \mathbb{R}_+)$  and  $(y, s) \in \Omega_m^-$  with  $y$  bounded as  $\epsilon_N \rightarrow 0$ .*

**6.3. Accuracy of the parametrices for  $\mathbf{Q}_m(w)$ . Error analysis.** The error in approximating  $\mathbf{Q}_m(w)$  by the parametrix  $\dot{\mathbf{Q}}_m^\pm(w)$  is the matrix  $\mathbf{E}_m^\pm(w) := \mathbf{Q}_m(w) \dot{\mathbf{Q}}_m^\pm(w)^{-1}$ . Since the jump contour  $\partial U$  for  $\dot{\mathbf{Q}}_m^\pm(w)$  is a subset of the jump contour  $\Sigma$  of  $\mathbf{Q}_m(w)$ , and since  $\det(\dot{\mathbf{Q}}_m^\pm(w)) = 1$  for  $w \in \mathbb{C} \setminus \partial U$ , it is clear that  $\mathbf{E}_m^\pm(w)$  is, like  $\mathbf{Q}_m(w)$ , analytic for  $w \in \mathbb{C} \setminus \Sigma$ , and takes its boundary values on  $\Sigma$  in the same classical sense as we require of  $\mathbf{Q}_m(w)$ . Also,  $\mathbf{E}_m^\pm(w) \rightarrow \mathbb{I}$  as  $w \rightarrow \infty$  since this is true of both factors.

We now consider the jump conditions satisfied by  $\mathbf{E}_m^\pm(w)$  as a consequence of those known to be satisfied by  $\mathbf{Q}_m(w)$  from the statement of Riemann-Hilbert Problem 6.1 and those satisfied by  $\dot{\mathbf{Q}}_m^\pm(w)$  from either Riemann-Hilbert Problem 6.2 or 6.3. If  $\xi \in \mathbb{R}_+$ , then

$$\begin{aligned} \mathbf{E}_{m+}^\pm(\xi) &= \mathbf{Q}_{m+}(\xi) \dot{\mathbf{Q}}_{m+}^\pm(\xi)^{-1} \\ &= \sigma_2 \mathbf{Q}_{m-}(\xi) \sigma_2 (\mathbb{I} + \mathbf{X}_m(\xi)) \sigma_2 \dot{\mathbf{Q}}_{m-}^\pm(\xi)^{-1} \sigma_2 \\ &= \sigma_2 \mathbf{E}_{m-}^\pm(\xi) \sigma_2 \cdot \left[ \sigma_2 \dot{\mathbf{Q}}_{m-}^\pm(\xi) \sigma_2 (\mathbb{I} + \mathbf{X}_m(\xi)) \sigma_2 \dot{\mathbf{Q}}_{m-}^\pm(\xi)^{-1} \sigma_2 \right]. \end{aligned} \quad (6.63)$$

If we write the matrix in brackets above as  $\mathbb{I} + \tilde{\mathbf{X}}_m^\pm(\xi)$ , we see that  $\tilde{\mathbf{X}}_m^\pm(\xi) = \mathbf{0}$  wherever  $\mathbf{X}_m(\xi) = \mathbf{0}$  (that is, everywhere except in the interval  $I$  near  $\xi = 1$ ). Using Proposition 6.2 and the fact that  $\mathbf{X}_m(\xi) = \mathcal{O}(\epsilon_N)$  for  $\xi \in I$ , we see that under the condition that  $(y, s) \in \Omega_m^+$  with  $y$  bounded we also have  $\tilde{\mathbf{X}}_m^\pm(\xi) = \mathcal{O}(\epsilon_N)$  for  $\xi \in I$ . Similarly, Proposition 6.4 shows that under the condition that  $(y, s) \in \Omega_m^-$  with  $y$  bounded the same estimate holds for  $\tilde{\mathbf{X}}_m^\pm(\xi)$ . If  $\xi \in \Sigma \setminus (\mathbb{R}_+ \cup \partial U)$ , then the parametrix has no discontinuity, so

$$\mathbf{E}_{m+}^\pm(\xi) = \mathbf{Q}_{m+}(\xi) \dot{\mathbf{Q}}_m^\pm(\xi)^{-1} = \mathbf{E}_{m-}^\pm(\xi) \dot{\mathbf{Q}}_m^\pm(\xi) (\mathbb{I} + \mathcal{O}(\epsilon_N)) \dot{\mathbf{Q}}_m^\pm(\xi)^{-1} \quad (6.64)$$

Therefore according to Proposition 6.2, if  $(y, s) \in \Omega_m^+$  with  $y$  bounded, we will have  $\mathbf{E}_{m+}^+(\xi) = \mathbf{E}_{m-}^+(\xi)(\mathbb{I} + \mathcal{O}(\epsilon_N))$  for such  $\xi$ . Analogously, Proposition 6.4 shows that  $(y, s) \in \Omega_m^-$  with  $y$  bounded implies that  $\mathbf{E}_{m+}^-(\xi) = \mathbf{E}_{m-}^-(\xi)(\mathbb{I} + \mathcal{O}(\epsilon_N))$  for such  $\xi$ . Finally, if  $\xi \in \partial U$ , then

$$\begin{aligned} \mathbf{E}_{m+}^+(\xi) &= \mathbf{Q}_{m+}(\xi) \dot{\mathbf{Q}}_{m+}^+(\xi)^{-1} \\ &= \mathbf{Q}_{m-}(\xi) \mathbf{V}_{\mathbf{Q}_m}(\xi) \begin{bmatrix} 1 & Y_m^+(\xi) \\ 0 & 1 \end{bmatrix}^{-1} \dot{\mathbf{Q}}_{m-}^+(\xi)^{-1} \\ &= \mathbf{E}_{m-}^+(\xi) \left( \dot{\mathbf{Q}}_{m-}^+(\xi) \mathbf{V}_{\mathbf{Q}_m}(\xi) \begin{bmatrix} 1 & Y_m^+(\xi) \\ 0 & 1 \end{bmatrix}^{-1} \dot{\mathbf{Q}}_{m-}^+(\xi)^{-1} \right). \end{aligned} \quad (6.65)$$

If  $(y, s) \in \Omega_m^+$  with  $y$  bounded, then according to Propositions 6.1 and 6.2 we find  $\mathbf{E}_{m+}^+(\xi) = \mathbf{E}_{m-}^+(\xi)(\mathbb{I} + \mathfrak{o}(1))$  where the size of the error term is  $\mathcal{O}(\epsilon_N^{1/6})$  at the worst (according to Corollary 6.1) but is more typically smaller according to the more precise estimates (6.21)–(6.23) enumerated in Proposition 6.1. Similarly, for  $\xi \in \partial U$ ,

$$\begin{aligned} \mathbf{E}_{m+}^-(\xi) &= \mathbf{Q}_{m+}(\xi) \dot{\mathbf{Q}}_{m+}^-(\xi)^{-1} \\ &= \mathbf{Q}_{m-}(\xi) \mathbf{V}_{\mathbf{Q}_m}(\xi) \begin{bmatrix} 1 & 0 \\ Y_m^-(\xi) & 1 \end{bmatrix}^{-1} \dot{\mathbf{Q}}_{m-}^-(\xi)^{-1} \\ &= \mathbf{E}_{m-}^-(\xi) \left( \dot{\mathbf{Q}}_{m-}^-(\xi) \mathbf{V}_{\mathbf{Q}_m}(\xi) \begin{bmatrix} 1 & 0 \\ Y_m^-(\xi) & 1 \end{bmatrix}^{-1} \dot{\mathbf{Q}}_{m-}^-(\xi)^{-1} \right), \end{aligned} \quad (6.66)$$

so if  $(y, s) \in \Omega_m^-$  with  $y$  bounded, then from Propositions 6.3 and 6.4 we find  $\mathbf{E}_{m+}^-(\xi) = \mathbf{E}_{m-}^-(\xi)(\mathbb{I} + \mathfrak{o}(1))$  with a similar characterization of the error term.

The dominant terms in the deviation of the jump matrix from the identity come from  $\xi \in \partial U$ . Therefore, we observe that the matrix  $\mathbf{E}_m^\pm$  satisfies the conditions of the following type of Riemann-Hilbert problem.

**Riemann-Hilbert Problem 6.4** (Small-Norm Problem for the Error). *Let  $m$  be an integer with  $|m| \leq B$ , and suppose that  $(y, s) \in \Omega_m^\pm$  with  $y$  bounded. Seek a matrix  $\mathbf{E}_m^\pm(w)$  with the following properties:*

**Analyticity:**  $\mathbf{E}_m^\pm(w)$  is analytic for  $w \in \mathbb{C} \setminus \Sigma$ , and is uniformly Hölder continuous up to the boundary of its domain of analyticity.

**Jump conditions:** The boundary values taken by  $\mathbf{E}_m^\pm(w)$  on the contour  $\Sigma$  are related as follows:

$$\mathbf{E}_{m+}^\pm(\xi) = \sigma_2 \mathbf{E}_{m-}^\pm(\xi) \sigma_2 \mathbf{V}_{\mathbf{E}_m^\pm}(\xi), \quad \xi > 0, \quad (6.67)$$

and

$$\mathbf{E}_{m+}^\pm(\xi) = \mathbf{E}_{m-}^\pm(\xi) \mathbf{V}_{\mathbf{E}_m^\pm}(\xi), \quad \xi \in \Sigma \setminus \mathbb{R}_+, \quad (6.68)$$

where the following estimates hold for the jump matrix. Firstly, we have  $\mathbf{V}_{\mathbf{E}_m^\pm}(\xi) = \mathbb{I}$  for  $\xi > 0$  with  $\log(|\xi|)$  sufficiently large, so  $\mathbf{V}_{\mathbf{E}_m^\pm} - \mathbb{I}$  is compactly supported and vanishes identically in a neighborhood of  $\xi = 0$ . Next, we have  $\|\mathbf{V}_{\mathbf{E}_m^\pm} - \mathbb{I}\|_{L^\infty(\Sigma)} = \mathcal{O}(\epsilon_N^{1/6})$ , or, more precisely,  $\|\mathbf{V}_{\mathbf{E}_m^\pm} - \mathbb{I}\|_{L^\infty(\Sigma)} = \mathcal{O}(e_m^\pm(y, s; \epsilon_N))$ , where for  $(y, s) \in \Omega_m^+$ ,

$$e_m^+(y, s; \epsilon_N) := \begin{cases} \epsilon_N^{1/3} e^{p_m} |y - \mathcal{Z}(\mathcal{U}_m)|, & p_m > 0 \text{ and } |y - \mathcal{Z}(\mathcal{U}_m)| \geq \delta e^{-p_m} \\ \epsilon_N^{1/3} |y - \mathcal{P}(\mathcal{U}_m)|^{-1}, & |y - \mathcal{P}(\mathcal{U}_m)| \leq \delta \\ \epsilon_N^{1/3}, & \text{elsewhere in } \Omega_m^+, \end{cases} \quad (6.69)$$

and for  $(y, s) \in \Omega_m^-$ ,

$$e_m^-(y, s; \epsilon_N) := \begin{cases} \epsilon_N^{1/3} e^{-p_{m-1}} |y - \mathcal{Z}(\mathcal{V}_m)|, & p_{m-1} < 0 \text{ and } |y - \mathcal{Z}(\mathcal{V}_m)| \geq \delta e^{p_{m-1}} \\ \epsilon_N^{1/3} |y - \mathcal{P}(\mathcal{V}_m)|^{-1}, & |y - \mathcal{P}(\mathcal{V}_m)| \leq \delta \\ \epsilon_N^{1/3}, & \text{elsewhere in } \Omega_m^-. \end{cases} \quad (6.70)$$

**Normalization:** the matrix  $\mathbf{E}_m^\pm(w)$  satisfies the condition

$$\lim_{w \rightarrow \infty} \mathbf{E}_m^\pm(w) = \mathbb{I}. \quad (6.71)$$

**Proposition 6.5.** *Suppose that  $m$  is an integer with  $|m| \leq B$ , and that  $(y, s) \in \Omega_m^\pm$  with  $y$  bounded. Then for sufficiently small  $\epsilon_N$ , there exists a unique solution  $\mathbf{E}_m^\pm(w)$  of Riemann-Hilbert Problem 6.4. The solution has expansions for small and large  $w$  of the form*

$$\mathbf{E}_m^\pm(w) = \begin{cases} [{}^0\mathbf{E}_m^\pm] + [{}^1\mathbf{E}_m^\pm](-w)^{1/2} + \mathcal{O}(w), & w \rightarrow 0 \\ \mathbb{I} + [{}^\infty\mathbf{E}_m^\pm](-w)^{-1/2} + \mathcal{O}(w^{-1}), & w \rightarrow \infty, \end{cases} \quad (6.72)$$

and we have the estimates

$$[{}^0\mathbf{E}_m^\pm] = \mathbb{I} + \mathcal{O}(e_m^\pm(y, s; \epsilon_N)), \quad [{}^1\mathbf{E}_m^\pm] = \mathcal{O}(e_m^\pm(y, s; \epsilon_N)), \quad \text{and} \quad [{}^\infty\mathbf{E}_m^\pm] = \mathcal{O}(e_m^\pm(y, s; \epsilon_N)) \quad (6.73)$$

holding uniformly for  $(y, s) \in \Omega_m^\pm$  with  $y$  bounded, where  $e_m^\pm(y, s; \epsilon_N)$  and  $e_m^\pm(y, s; \epsilon_N)$  are defined by (6.69) and (6.70) respectively.

*Proof.* By the substitution  $\mathbf{F}_m^\pm(\tau) := \mathbf{E}_m^\pm(\tau^2)$  for  $\Im\{\tau\} > 0$  and  $\mathbf{F}_m^\pm(\tau) := \sigma_2 \mathbf{E}_m^\pm(\tau^2) \sigma_2$  for  $\Im\{\tau\} < 0$ , we translate the conditions of Riemann-Hilbert Problem 6.4 into a list of conditions satisfied by  $\mathbf{F}_m^\pm(\tau)$ . This “unfolding” of  $w$  to the  $\tau$ -plane implies that  $\mathbf{F}_m^\pm(\tau)$  is analytic on the complement of a compact contour  $\Sigma'$ , along which it satisfies jump conditions of the form  $\mathbf{F}_{m+}^\pm = \mathbf{F}_{m-}^\pm \mathbf{V}_{\mathbf{F}_m^\pm}$  where  $\mathbf{V}_{\mathbf{F}_m^\pm} - \mathbb{I}$  satisfies the uniform estimate  $\|\mathbf{V}_{\mathbf{F}_m^\pm} - \mathbb{I}\|_{L^\infty(\Sigma')} = \mathcal{O}(e_m^\pm(y, s; \epsilon_N))$  and  $e_m^\pm(y, s; \epsilon_N)$  is defined by (6.69) for  $(y, s) \in \Omega_m^+$  and by (6.70) for  $(y, s) \in \Omega_m^-$ . Since these estimates are all dominated by  $\mathcal{O}(\epsilon_N^{1/6})$ , for sufficiently small  $\epsilon_N$ , the small-norm theory of matrix Riemann-Hilbert problems in, say, the  $L^2$ -sense implies that  $\mathbf{F}_m^\pm(\tau)$  may be constructed by iteration applied to a suitable singular integral equation, and it follows from that theory that an estimate of the form  $\|\mathbf{F}_m^\pm - \mathbb{I}\|_{L^\infty(K)} = \mathcal{O}(e_m^\pm(y, s; \epsilon_N))$  actually holds whenever  $K$  is a compact set disjoint from the jump contour  $\Sigma'$ . In particular,  $\mathbf{F}_m^\pm(\tau)$  is analytic on some compact neighborhood  $K$  of the origin  $\tau = 0$ , so it has a Taylor expansion that we write in the form

$$\mathbf{F}_m^\pm(\tau) = [{}^0\mathbf{F}_m^\pm] + [{}^1\mathbf{F}_m^\pm]\tau + \mathcal{O}(\tau^2), \quad \tau \rightarrow 0, \quad (6.74)$$

and by the uniform estimate on  $\mathbf{F}_m^\pm(\tau) - \mathbb{I}$  we have  $[{}^0\mathbf{F}_m^\pm] = [{}^0\mathbf{E}_m^\pm] = \mathbb{I} + \mathcal{O}(\|\mathbf{F}_m^\pm - \mathbb{I}\|_{L^\infty(K)})$ . Applying the Cauchy integral formula to  $\mathbf{F}_m^\pm(\tau) - \mathbb{I}$  on a small loop in  $K$  surrounding the origin then shows that  $[{}^1\mathbf{F}_m^\pm] = i[{}^1\mathbf{E}_m^\pm] = \mathcal{O}(\|\mathbf{F}_m^\pm - \mathbb{I}\|_{L^\infty(K)})$ . Also,  $\mathbf{F}_m^\pm(\tau)$  tends to the identity as  $\tau \rightarrow \infty$  and it is analytic for sufficiently large  $|\tau|$  so it has a Laurent expansion of the form

$$\mathbf{F}_m^\pm(\tau) = \mathbb{I} + [{}^\infty\mathbf{F}_m^\pm]\tau^{-1} + \mathcal{O}(\tau^{-2}), \quad \tau \rightarrow \infty, \quad (6.75)$$

Now we can take  $K$  as an annulus enclosing a large circle. Applying the Cauchy integral formula on such a circle to  $\mathbf{F}_m^\pm(\tau) - \mathbb{I}$  we find that  $[{}^\infty\mathbf{F}_m^\pm] = -i[{}^\infty\mathbf{E}_m^\pm] = \mathcal{O}(\|\mathbf{F}_m^\pm - \mathbb{I}\|_{L^\infty(K)})$ . This completes the proof.  $\square$

The magnitude of the expansion coefficients of  $\mathbf{E}_m^+(w) \approx \mathbb{I}$  and of  $\mathbf{E}_{m+1}^-(w) \approx \mathbb{I}$  in the interlocking regions  $\Omega_m^+$  and  $\Omega_{m+1}^-$  respectively is illustrated in Figure 6.5. In particular it is clear that the order of accuracy is  $\mathcal{O}(\epsilon_N^{1/3})$  except very close to the common boundaries between teeth and notches. Moreover, the order of accuracy is worst right along these common boundaries where it is  $\mathcal{O}(\epsilon_N^{1/3} e^{|p_m|/2})$  which in turn is largest just at the tips of the teeth ( $|p_m| = \frac{1}{3} \log(\epsilon_N^{-1})$ ), where the estimate reduces to  $\mathcal{O}(\epsilon_N^{1/6})$ .

## 7. PROOFS OF THEOREMS

**7.1. Exact formulae for the potentials.** Beginning with the matrix  $\mathbf{N}(w) = \mathbf{N}_N^g(w; x, t)$  we may now express  $\mathbf{N}(w)$  in terms of explicitly known quantities and the error matrix  $\mathbf{E}_m^\pm(w)$  characterized in Proposition 6.5. For example, if  $|w|$  is sufficiently small then for  $\Im\{w\} > 0$  we have

$$\mathbf{N}(w) = \mathbf{O}(w) = \mathbf{P}(w) i \sigma_1 = \mathbf{Q}_m(w) \dot{\mathbf{P}}_m^{\text{out}}(w) i \sigma_1 = \mathbf{E}_m^\pm(w) \dot{\mathbf{Q}}_m^\pm(w) \dot{\mathbf{P}}_m^{\text{out}}(w) i \sigma_1, \quad (7.1)$$

while for  $\Im\{w\} < 0$  we have the same formula just replacing  $\mathbf{P}(w)$  and  $\dot{\mathbf{P}}_m^{\text{out}}(w)$  by  $-\mathbf{P}(w)$  and  $-\dot{\mathbf{P}}_m^{\text{out}}(w)$  respectively. But from (3.7) we see that

$$\text{sgn}(\Im\{w\}) \dot{\mathbf{P}}_m^{\text{out}}(w) = \left( \mathbb{I} + \frac{1-2m}{S} \sigma_3 (-w)^{1/2} + \mathcal{O}(w) \right) (-1)^{m+1} i \sigma_2, \quad w \rightarrow 0, \quad (7.2)$$

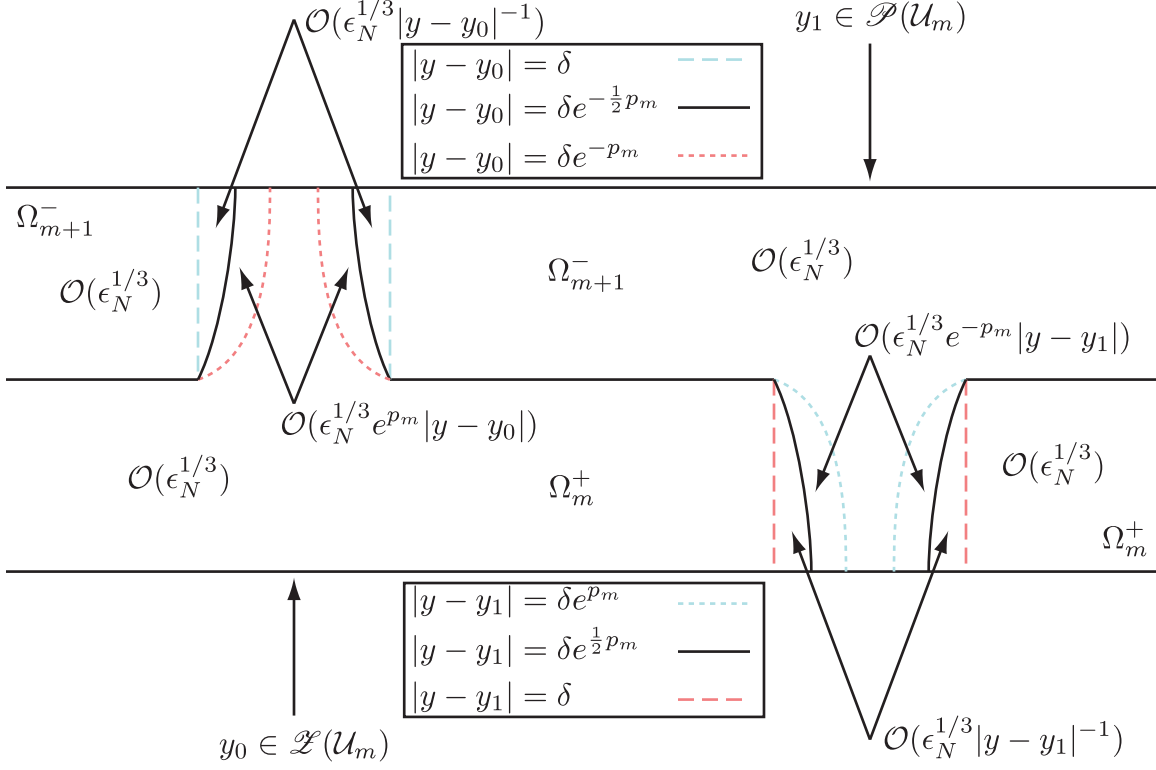


FIGURE 6.5. Estimates  $e_m^+(y, s; \epsilon_N)$  for expansion coefficients of  $\mathbf{E}_m^+(w) - \mathbb{I}$  with  $(y, s) \in \Omega_m^+$  and  $e_{m+1}^-(y, s; \epsilon_N)$  for expansion coefficients of  $\mathbf{E}_{m+1}^-(w) - \mathbb{I}$  with  $(y, s) \in \Omega_{m+1}^-$ . The boundaries of the two interlocking regions are drawn with solid black curves.

where  $S := (-w_*)^{1/2} > 0$ . Therefore, in terms of the expansion (6.72) of  $\mathbf{E}_m^\pm(w)$  for small  $|w|$ , and the corresponding expansion of  $\dot{\mathbf{Q}}_m^\pm(w)$ :

$$\dot{\mathbf{Q}}_m^\pm(w) = [{}^0_0 \dot{\mathbf{Q}}_m^\pm] + [{}^0_1 \dot{\mathbf{Q}}_m^\pm](-w)^{1/2} + \mathcal{O}(w), \quad w \rightarrow 0, \quad (7.3)$$

we find the following formulae for certain expansion coefficients of  $\mathbf{N}(w)$  (see (1.26)):

$$[{}^0_0 \mathbf{N}_N^g](x, t) = [{}^0_0 \mathbf{E}_m^\pm][{}^0_0 \dot{\mathbf{Q}}_m^\pm](-1)^{m+1} i \sigma_2, \quad (7.4)$$

and

$$[{}^0_0 \mathbf{N}_N^g](x, t)^{-1} [{}^0_1 \mathbf{N}_N^g](x, t) = \sigma_2 [{}^0_0 \dot{\mathbf{Q}}_m^\pm]^{-1} [{}^0_0 \mathbf{E}_m^\pm]^{-1} [{}^0_1 \mathbf{E}_m^\pm][{}^0_0 \dot{\mathbf{Q}}_m^\pm] \sigma_2 + \sigma_2 [{}^0_0 \dot{\mathbf{Q}}_m^\pm]^{-1} [{}^0_1 \dot{\mathbf{Q}}_m^\pm] \sigma_2 - \frac{1-2m}{S} \sigma_3. \quad (7.5)$$

Similarly, if  $|w|$  is sufficiently large, then we have

$$\mathbf{N}(w) = \mathbf{O}(w) = \mathbf{P}(w) = \mathbf{Q}_m(w) \dot{\mathbf{P}}_m^{\text{out}}(w) = \mathbf{E}_m^\pm(w) \dot{\mathbf{Q}}_m^\pm(w) \dot{\mathbf{P}}_m^{\text{out}}(w), \quad (7.6)$$

and from (3.7) we have

$$\dot{\mathbf{P}}_m^{\text{out}}(w) = \mathbb{I} + (1-2m)S\sigma_3(-w)^{-1/2} + \mathcal{O}(w^{-1}), \quad w \rightarrow \infty, \quad (7.7)$$

so in terms of the expansion (6.72) of  $\mathbf{E}_m^\pm(w)$  for large  $|w|$ , and the corresponding expansion of  $\dot{\mathbf{Q}}_m^\pm(w)$ :

$$\dot{\mathbf{Q}}_m^\pm(w) = \mathbb{I} + [{}^\infty_1 \dot{\mathbf{Q}}_m^\pm](-w)^{-1/2} + \mathcal{O}(w^{-1}), \quad w \rightarrow \infty, \quad (7.8)$$

we obtain an exact formula for the coefficient  $[{}^\infty_1 \mathbf{N}_N^g](x, t)$  (see (1.25)):

$$[{}^\infty_1 \mathbf{N}_N^g](x, t) = [{}^\infty_1 \mathbf{E}_m^\pm] + [{}^\infty_1 \dot{\mathbf{Q}}_m^\pm] + (1-2m)S\sigma_3. \quad (7.9)$$

Combining (7.4) with (1.27) from Definition 1.1 we obtain the exact formulae:

$$\cos(\tfrac{1}{2}u_N(x, t)) = (-1)^m [{}^0_0 \mathbf{E}_m^\pm][{}^0_0 \dot{\mathbf{Q}}_m^\pm]_{12} \quad \text{and} \quad \sin(\tfrac{1}{2}u_N(x, t)) = (-1)^m [{}^0_0 \mathbf{E}_m^\pm][{}^0_0 \dot{\mathbf{Q}}_m^\pm]_{22}. \quad (7.10)$$

Similarly, combining (7.5) and (7.9) with (1.28) gives

$$\epsilon_N \frac{\partial u_N}{\partial t}(x, t) = [\infty \mathbf{E}_m^\pm]_{12} + [\infty \dot{\mathbf{Q}}_m^\pm]_{12} - [\begin{smallmatrix} 0 \\ 0 \end{smallmatrix} \dot{\mathbf{Q}}_m^\pm]^{-1} [\begin{smallmatrix} 0 \\ 0 \end{smallmatrix} \mathbf{E}_m^\pm]^{-1} [\begin{smallmatrix} 0 \\ 1 \end{smallmatrix} \mathbf{E}_m^\pm] [\begin{smallmatrix} 0 \\ 0 \end{smallmatrix} \dot{\mathbf{Q}}_m^\pm]_{21} - [\begin{smallmatrix} 0 \\ 0 \end{smallmatrix} \dot{\mathbf{Q}}_m^\pm]^{-1} [\begin{smallmatrix} 0 \\ 1 \end{smallmatrix} \dot{\mathbf{Q}}_m^\pm]_{21}. \quad (7.11)$$

Now suppose that  $(x, t)$  are such that  $(y, s) = (\epsilon^{-2/3} r(x, t), s(x, t)) \in \Omega_m^\pm$  with  $y$  bounded. Then we can apply Proposition 6.5 and the boundedness of  $[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix} \dot{\mathbf{Q}}_m^\pm] = \dot{\mathbf{Q}}_m^\pm(0)$  and its inverse as guaranteed by Propositions 6.2 and 6.4 to obtain:

$$\begin{aligned} \cos(\tfrac{1}{2} u_N(x, t)) &= \dot{C}_{N,m}^\pm(x, t) + \mathcal{O}(e_m^\pm(y, s; \epsilon_N)) \\ \sin(\tfrac{1}{2} u_N(x, t)) &= \dot{S}_{N,m}^\pm(x, t) + \mathcal{O}(e_m^\pm(y, s; \epsilon_N)) \\ \epsilon_N \frac{\partial u_N}{\partial t}(x, t) &= \dot{G}_{N,m}^\pm(x, t) + \mathcal{O}(e_m^\pm(y, s; \epsilon_N)), \end{aligned} \quad (7.12)$$

all holding for  $(y, s) \in \Omega_m^\pm$  with  $y$  bounded, where

$$\begin{aligned} \dot{C}_{N,m}^\pm(x, t) &:= (-1)^m [\begin{smallmatrix} 0 \\ 0 \end{smallmatrix} \dot{\mathbf{Q}}_m^\pm]_{12} \\ \dot{S}_{N,m}^\pm(x, t) &:= (-1)^m [\begin{smallmatrix} 0 \\ 0 \end{smallmatrix} \dot{\mathbf{Q}}_m^\pm]_{22} \\ \dot{G}_{N,m}^\pm(x, t) &:= [\infty \dot{\mathbf{Q}}_m^\pm]_{12} - [\begin{smallmatrix} 0 \\ 0 \end{smallmatrix} \dot{\mathbf{Q}}_m^\pm]^{-1} [\begin{smallmatrix} 0 \\ 1 \end{smallmatrix} \dot{\mathbf{Q}}_m^\pm]_{21} = [\infty \dot{\mathbf{Q}}_m^\pm]_{12} + [\begin{smallmatrix} 0 \\ 0 \end{smallmatrix} \dot{\mathbf{Q}}_m^\pm]_{21} [\begin{smallmatrix} 0 \\ 1 \end{smallmatrix} \dot{\mathbf{Q}}_m^\pm]_{11} - [\begin{smallmatrix} 0 \\ 0 \end{smallmatrix} \dot{\mathbf{Q}}_m^\pm]_{11} [\begin{smallmatrix} 0 \\ 1 \end{smallmatrix} \dot{\mathbf{Q}}_m^\pm]_{21}. \end{aligned} \quad (7.13)$$

(In the last equality we used the fact that  $[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix} \dot{\mathbf{Q}}_m^\pm] = \dot{\mathbf{Q}}_m^\pm(0)$  has determinant one.) Since  $\dot{\mathbf{Q}}_m^\pm(w) = \mathbf{F}_m^\pm(w) = \mathbf{G}_m^\pm(i(-w)^{1/2})$ , the necessary expansion coefficients of  $\dot{\mathbf{Q}}_m^\pm(w)$  (see (7.3) and (7.8)) can be obtained directly from the partial-fractions expansions (6.37) and (6.58). In this way, we get

$$\begin{aligned} [\begin{smallmatrix} 0 \\ 0 \end{smallmatrix} \dot{\mathbf{Q}}_m^+] &= \mathbb{I} + \frac{1}{w_*} (i\sigma_2[\mathbf{b}_m^+ - iS\mathbf{d}_m^+], \mathbf{b}_m^+ - iS\mathbf{d}_m^+) \\ [\begin{smallmatrix} 0 \\ 1 \end{smallmatrix} \dot{\mathbf{Q}}_m^+] &= \frac{1}{w_* S} (-i\sigma_2[2\mathbf{b}_m^+ - iS\mathbf{d}_m^+], 2\mathbf{b}_m^+ - iS\mathbf{d}_m^+) \\ [\infty \dot{\mathbf{Q}}_m^+] &= -i(-i\sigma_2\mathbf{d}_m^+, \mathbf{d}_m^+), \end{aligned} \quad (7.14)$$

implying that

$$\begin{aligned} \dot{C}_{N,m}^+(x, t) &= (-1)^m \left[ \frac{b_{m,1}^+}{w_*} + \frac{id_{m,1}^+}{S} \right] \\ \dot{S}_{N,m}^+(x, t) &= (-1)^m \left[ 1 + \frac{b_{m,2}^+}{w_*} + \frac{id_{m,2}^+}{S} \right] \\ \dot{G}_{N,m}^+(x, t) &= \frac{1}{w_*^2} [iw_*(1 - w_*)d_{m,1}^+ + 2Sb_{m,1}^+ + ib_{m,1}^+d_{m,2}^+ - ib_{m,2}^+d_{m,1}^+], \end{aligned} \quad (7.15)$$

where the components of the vectors  $\mathbf{b}_m^+$  and  $\mathbf{d}_m^+$  are given by (6.40) with (6.41) and we recall that  $S := (-w_*)^{1/2} > 0$ . In the same way one sees that

$$\begin{aligned} [\begin{smallmatrix} 0 \\ 0 \end{smallmatrix} \dot{\mathbf{Q}}_m^-] &= \mathbb{I} + \frac{1}{w_*} (\mathbf{a}_m^- - iS\mathbf{c}_m^-, -i\sigma_2[\mathbf{a}_m^- - iS\mathbf{c}_m^-]) \\ [\begin{smallmatrix} 0 \\ 1 \end{smallmatrix} \dot{\mathbf{Q}}_m^-] &= \frac{1}{w_* S} (2\mathbf{a}_m^- - iS\mathbf{c}_m^-, i\sigma_2[2\mathbf{a}_m^- - iS\mathbf{c}_m^-]) \\ [\infty \dot{\mathbf{Q}}_m^-] &= -i(\mathbf{c}_m^-, i\sigma_2\mathbf{c}_m^-), \end{aligned} \quad (7.16)$$

implying that

$$\begin{aligned} \dot{C}_{N,m}^-(x, t) &= (-1)^{m+1} \left[ \frac{a_{m,2}^-}{w_*} + \frac{ic_{m,2}^-}{S} \right] \\ \dot{S}_{N,m}^-(x, t) &= (-1)^m \left[ 1 + \frac{a_{m,1}^-}{w_*} + \frac{ic_{m,1}^-}{S} \right] \\ \dot{G}_{N,m}^-(x, t) &= \frac{1}{w_*^2} [iw_*(1 - w_*)c_{m,2}^- + 2Sa_{m,2}^- + ia_{m,2}^-c_{m,1}^- - ia_{m,1}^-c_{m,2}^-], \end{aligned} \quad (7.17)$$

where  $\mathbf{a}_m^-$  and  $\mathbf{c}_m^-$  are given by (6.61) with (6.62) and we have  $S := (-w_*)^{1/2} > 0$ .

Finally, substituting from (6.40)–(6.41) and (6.61)–(6.62) gives

$$\begin{aligned}\dot{C}_{N,m}^\pm(x,t) &= \pm(-1)^{m+1} \frac{2R_m^\pm}{[R_m^\pm]^2 + 1} \\ \dot{S}_{N,m}^\pm(x,t) &= (-1)^{m+1} \frac{[R_m^\pm]^2 - 1}{[R_m^\pm]^2 + 1} \\ \dot{G}_{N,m}^\pm(x,t) &= 16w_* \frac{[J_m^\pm]^2 K_m^\pm + 2S[J_m^\pm]^3 - w_*[J_m^\pm]^2 K_m^\pm + 32w_* S J_m^\pm - 16w_*^2 K_m^\pm + 16w_*^3 K_m^\pm}{([J_m^\pm]^2 + 16w_*^2)^2 + 64w_*^2 (J_m^\pm + S K_m^\pm)^2},\end{aligned}\tag{7.18}$$

where

$$R_m^\pm := \frac{-8w_*(J_m^\pm + S K_m^\pm)}{[J_m^\pm]^2 + 16w_*^2},\tag{7.19}$$

and where  $J_m^+$  and  $K_m^+$  are defined by (6.33) while  $J_m^-$  and  $K_m^-$  are defined by (6.54).

**7.2. Proof of Theorem 1.1 (Initial accuracy of fluxon condensates).** The fluxon condensate  $u_N(x,t)$  is an even function of  $x$ , so it suffices to consider  $x \approx x_{\text{crit}} > 0$ . The line segment  $t = 0$  with  $\Delta x = \mathcal{O}(\epsilon_N^{2/3})$  is mapped under  $(y, s) = (\epsilon^{-2/3}r(x,t), s(x,t))$  into the union  $\Omega_0^+ \cup \Omega_1^-$ , assuming that  $\epsilon_N$  sufficiently small. More specifically, it is mapped into the part of  $\Omega_0^+ \cup \Omega_1^-$  with  $y = \mathcal{O}(1)$  and  $p_0 = \mathcal{O}(\epsilon_N^{1/3})$ . Using the facts that  $\mathcal{U}_0(y) = \mathcal{V}_1(y) = 1$  and  $\mathcal{V}_0(y) = \mathcal{U}_1(y) = -y/6$ , and hence also  $H_0(y) = H_1(y) = y^2/24$ , it is easy to see from the definitions (6.33) and (6.54) (using also  $h_0^+(w_*) = h_1^-(w_*) = -4w_* W'(w_*)$ ) that for  $t = 0$  and  $\Delta x = \mathcal{O}(\epsilon_N^{2/3})$ ,

$$J_0^+ = \mathcal{O}(\epsilon_N^{1/3}) \quad \text{and} \quad J_1^- = \mathcal{O}(\epsilon_N^{1/3})\tag{7.20}$$

while

$$K_0^+ = -2S + \mathcal{O}(\epsilon_N^{1/3}) \quad \text{and} \quad K_1^- = -2S + \mathcal{O}(\epsilon_N^{1/3}),\tag{7.21}$$

where  $S := (-w_*)^{1/2} > 0$ . According to (7.19) we then have

$$R_0^+ = -1 + \mathcal{O}(\epsilon_N^{1/3}) \quad \text{and} \quad R_1^- = -1 + \mathcal{O}(\epsilon_N^{1/3}).\tag{7.22}$$

It follows from these facts and (7.18) that for such  $x$  and  $t = 0$ ,

$$\dot{C}_{N,0}^+(x,0) = 1 + \mathcal{O}(\epsilon_N^{1/3}) \quad \text{and} \quad \dot{C}_{N,1}^-(x,0) = 1 + \mathcal{O}(\epsilon_N^{1/3}),\tag{7.23}$$

$$\dot{S}_{N,0}^+(x,0) = \mathcal{O}(\epsilon_N^{1/3}) \quad \text{and} \quad \dot{S}_{N,1}^-(x,0) = \mathcal{O}(\epsilon_N^{1/3}),\tag{7.24}$$

and

$$\dot{G}_{N,0}^+(x,0) = -\left(S + \frac{1}{S}\right) + \mathcal{O}(\epsilon_N^{1/3}) \quad \text{and} \quad \dot{G}_{N,1}^-(x,0) = -\left(S + \frac{1}{S}\right) + \mathcal{O}(\epsilon_N^{1/3}).\tag{7.25}$$

Furthermore, since  $w_*$  is an analytic function of  $(x,t)$  with  $w_*(x_{\text{crit}}, 0) = -1$ , we have  $w_* = -1 + \mathcal{O}(\Delta x, t) = -1 + \mathcal{O}(\Delta x) = -1 + \mathcal{O}(\epsilon_N^{2/3})$ . Therefore, we may equivalently write the above formulae in the form

$$\dot{G}_{N,0}^+(x,0) = -2 + \mathcal{O}(\epsilon_N^{1/3}) \quad \text{and} \quad \dot{G}_{N,1}^-(x,0) = -2 + \mathcal{O}(\epsilon_N^{1/3}), \quad \Delta x = \mathcal{O}(\epsilon_N^{2/3}).\tag{7.26}$$

Also, since the function  $G(\cdot)$  in the initial data is differentiable at  $x_{\text{crit}}$  and  $G(x_{\text{crit}}) = -2$ , we can also write

$$\dot{G}_{N,0}^+(x,0) = G(x) + \mathcal{O}(\epsilon_N^{1/3}) \quad \text{and} \quad \dot{G}_{N,1}^-(x,0) = G(x) + \mathcal{O}(\epsilon_N^{1/3}), \quad \Delta x = \mathcal{O}(\epsilon_N^{2/3}).\tag{7.27}$$

Finally, recalling (7.12), the fact that  $e_0^+(y, s; \epsilon_N) = \mathcal{O}(\epsilon_N^{1/3})$  for  $(y, s) \in \Omega_0^+$  with  $y$  bounded and  $p_0 = \mathcal{O}(\epsilon_N^{1/3})$ , and the fact that  $e_1^-(y, s; \epsilon_N) = \mathcal{O}(\epsilon_N^{1/3})$  for  $(y, s) \in \Omega_1^-$  with  $y$  bounded and  $p_0 = \mathcal{O}(\epsilon_N^{1/3})$ , we arrive at

$$u_N(x,0) = \mathcal{O}(\epsilon_N^{1/3}) \pmod{4\pi} \quad \text{and} \quad \epsilon_N \frac{\partial u_N}{\partial t}(x,0) = G(x) + \mathcal{O}(\epsilon_N^{1/3}), \quad \Delta x = \mathcal{O}(\epsilon_N^{2/3}),\tag{7.28}$$

which completes the proof of the Theorem.

**7.3. Proof of Theorem 1.2 (Main approximation theorem).** Firstly, we wish to simplify the expressions for  $\dot{C}_{N,m}^\pm(x,t)$  and  $\dot{S}_{N,m}^\pm(x,t)$  given in (7.18) by localizing near the critical point  $\Delta x = t = 0$ , which allows us to write the coordinates  $r(x,t)$  and  $s(x,t)$  explicitly in terms of  $x$  and  $t$  at the cost of some small error terms according to (4.38). By analogy with the definition (6.11) of  $p_m$  for general  $m \in \mathbb{Z}$ , we define a shifted and scaled coordinate  $q_m$  equivalent to  $t$  by the relation

$$t = \frac{2m}{3}\epsilon_N \log(\epsilon_N^{-1}) + \epsilon_N q_m, \quad m \in \mathbb{Z}. \quad (7.29)$$

Recall the spatial coordinate  $z$  exactly proportional to  $\Delta x$  as defined by (1.43). We will now show that there is essentially no additional cost (beyond that already introduced in approximating  $\mathbf{E}_m^\pm(w)$ ) in explicitly introducing  $\Delta x$  and  $t$  as coordinates.

**Proposition 7.1.** *Suppose that  $(y, s) \in \Omega_m^+$  with  $y$  bounded and  $m$  fixed. Then*

$$\dot{C}_{N,m}^+(x,t) = \ddot{C}_{N,m}^+(x,t) + \mathcal{O}(e_m^+(y,s;\epsilon_N)) \quad \text{and} \quad \dot{S}_{N,m}^+(x,t) = \ddot{S}_{N,m}^+(x,t) + \mathcal{O}(e_m^+(y,s;\epsilon_N)) \quad (7.30)$$

where

$$\ddot{C}_{N,m}^+(x,t) := (-1)^{m+1} \frac{2\dot{R}_m^+}{[\dot{R}_m^+]^2 + 1} \quad \text{and} \quad \ddot{S}_{N,m}^+(x,t) := (-1)^{m+1} \frac{[\dot{R}_m^+]^2 - 1}{[\dot{R}_m^+]^2 + 1} \quad (7.31)$$

and where

$$\dot{R}_m^+ := \frac{-2^{4(1-m)}\nu^{-2m/3}e^{q_m}\mathcal{U}_m(z)}{2^{-8m}\nu^{-2(1+2m)/3}\epsilon_N^{2/3}e^{2q_m}(2H_m(z)\mathcal{U}_m(z) - \mathcal{U}_m'(z))^2 + 16}. \quad (7.32)$$

Similarly, suppose that  $(y, s) \in \Omega_m^-$  with  $y$  bounded and  $m$  fixed. Then

$$\dot{C}_{N,m}^-(x,t) = \ddot{C}_{N,m}^-(x,t) + \mathcal{O}(e_m^-(y,s;\epsilon_N)) \quad \text{and} \quad \dot{S}_{N,m}^-(x,t) = \ddot{S}_{N,m}^-(x,t) + \mathcal{O}(e_m^-(y,s;\epsilon_N)) \quad (7.33)$$

where

$$\ddot{C}_{N,m}^-(x,t) := (-1)^m \frac{2\dot{R}_m^-}{[\dot{R}_m^-]^2 + 1} \quad \text{and} \quad \ddot{S}_{N,m}^-(x,t) := (-1)^{m+1} \frac{[\dot{R}_m^-]^2 - 1}{[\dot{R}_m^-]^2 + 1} \quad (7.34)$$

and where

$$\dot{R}_m^- := \frac{-2^{4m}\nu^{2(m-1)/3}e^{-q_{m-1}}\mathcal{V}_m(z)}{2^{8(m-1)}\nu^{-2+4m/3}\epsilon_N^{2/3}e^{-2q_{m-1}}(2H_m(z)\mathcal{V}_m(z) - \mathcal{V}_m'(z))^2 + 16}. \quad (7.35)$$

*Proof.* First, we localize near criticality by isolating the terms in  $R_m^\pm$  depending “slowly” on  $x$  and  $t$  and replacing  $x$  with  $x_{\text{crit}}$  and  $t$  with zero in these terms. The “slowly varying” quantities are  $w_* = w_*(x,t)$  and the following:

$$\begin{aligned} \alpha_m^\pm &:= \frac{h_m^\pm(w_*)}{4w_*W'(w_*)^2} \\ \beta_m^\pm &:= \frac{4w_*h_m^\pm(w_*)}{W'(w_*)} \\ \gamma_m^\pm &:= \frac{4w_*h_m^{\pm'}(w_*)W'(w_*) - 4w_*h_m^\pm(w_*)W''(w_*) - 4h_m^\pm(w_*)W'(w_*)}{W'(w_*)^3}. \end{aligned} \quad (7.36)$$

Recalling the definitions of  $J_m^\pm$  and  $K_m^\pm$  given by (6.33) and (6.54), one obtains from (7.19) the exact expressions

$$\begin{aligned} R_m^+ &= \frac{\beta_m^+ e^{p_m} \mathcal{U}_m(y) + \gamma_m^+ \epsilon_N^{1/3} e^{p_m} B_{m,12}(y)}{[\alpha_m^+ \epsilon_N^{1/3} e^{p_m} B_{m,12}(y)]^2 + 16w_*^2}, \quad \text{where} \quad B_{m,12}(y) = 2H_m(y)\mathcal{U}_m(y) - \mathcal{U}_m'(y) \\ R_m^- &= \frac{\beta_m^- e^{-p_{m-1}} \mathcal{V}_m(y) + \gamma_m^- \epsilon_N^{1/3} e^{-p_{m-1}} B_{m,21}(y)}{[\alpha_m^- \epsilon_N^{1/3} e^{-p_{m-1}} B_{m,21}(y)]^2 + 16w_*^2}, \quad \text{where} \quad B_{m,21}(y) = \mathcal{V}_m'(y) - 2H_m(y)\mathcal{V}_m(y). \end{aligned} \quad (7.37)$$



Now, for each  $m \in \mathbb{Z}$ ,  $\alpha_m^\pm$ ,  $\beta_m^\pm$ ,  $\gamma_m^\pm$ , and  $w_*$  are all independent of  $\epsilon_N$ , depending only on  $(x, t)$ . Moreover, they are all analytic functions of  $(x, t)$  near criticality. Computing their limits at criticality using (4.55) and (4.61) together with (6.18) and (6.42) we obtain:

$$\begin{aligned}\alpha_m^\pm &= -2^{-2\pm(2-4m)}\nu^{(-2\pm(1-2m))/3} + \mathcal{O}(\Delta x, t) \\ \beta_m^\pm &= -2^{2\pm(2-4m)}\nu^{(-1\pm(1-2m))/3} + \mathcal{O}(\Delta x, t) \\ \gamma_m^\pm &= \mathcal{O}(\Delta x, t),\end{aligned}\tag{7.38}$$

and for  $w_*$  we have the corresponding expression (4.56). Since by assumption  $m$  and  $y$  are bounded, it follows immediately that  $s(x, t) = \mathcal{O}(\epsilon_N \log(\epsilon_N^{-1}))$  and that  $r(x, t) = \mathcal{O}(\epsilon_N^{2/3})$ . Then, since the Jacobian matrix  $\partial(r, s)/\partial(x, t)$  is diagonal and invertible at criticality according to (4.38), it also follows that  $\Delta x = \mathcal{O}(\epsilon_N^{2/3})$  and  $t = \mathcal{O}(\epsilon_N \log(\epsilon_N^{-1}))$ , and consequently all error terms of the form  $\mathcal{O}(\Delta x, t)$  may be replaced with  $\mathcal{O}(\epsilon_N^{2/3})$ . Therefore,

$$\begin{aligned}R_m^+ &= \frac{-2^{4(1-m)}\nu^{-2m/3}e^{p_m}\mathcal{U}_m(y)(1 + \mathcal{O}(\epsilon_N^{2/3})) + \epsilon_N^{1/3}e^{p_m}B_{m,12}(y)\mathcal{O}(\epsilon_N^{2/3})}{[-2^{-4m}\nu^{-(1+2m)/3}\epsilon_N^{1/3}e^{p_m}B_{m,12}(y)(1 + \mathcal{O}(\epsilon_N^{2/3}))]^2 + 16(1 + \mathcal{O}(\epsilon_N^{2/3}))} \\ R_m^- &= \frac{-2^{4m}\nu^{2(m-1)/3}e^{-p_{m-1}}\mathcal{V}_m(y)(1 + \mathcal{O}(\epsilon_N^{2/3})) + \epsilon_N^{1/3}e^{-p_{m-1}}B_{m,21}(y)\mathcal{O}(\epsilon_N^{2/3})}{[-2^{4(m-1)}\nu^{-1+2m/3}\epsilon_N^{1/3}e^{-p_{m-1}}B_{m,21}(y)(1 + \mathcal{O}(\epsilon_N^{2/3}))]^2 + 16(1 + \mathcal{O}(\epsilon_N^{2/3}))}.\end{aligned}\tag{7.39}$$

The denominators are bounded away from zero and, according to Proposition 5.2,  $B_{m,12}(y)$  and  $B_{m,21}(y)$  have simple poles at the points of  $\mathcal{P}(\mathcal{U}_m) = \mathcal{P}(\mathcal{V}_m)$  implying that

$$\begin{aligned}\epsilon_N^{1/3}e^{p_m}B_{m,12}(y) &= \mathcal{O}(1), \quad \text{for } (y, s) \in \Omega_m^+ \\ \epsilon_N^{1/3}e^{-p_{m-1}}B_{m,21}(y) &= \mathcal{O}(1), \quad \text{for } (y, s) \in \Omega_m^-.\end{aligned}\tag{7.40}$$

Therefore, the formulae (7.39) can be written in the form

$$\begin{aligned}R_m^+ &= \frac{-2^{4(1-m)}\nu^{-2m/3}e^{p_m}\mathcal{U}_m(y)}{2^{-8m}\nu^{-2(1+2m)/3}\epsilon_N^{2/3}e^{2p_m}B_{m,12}(y)^2 + 16} + \mathcal{O}(\epsilon_N^{2/3}e^{p_m}\mathcal{U}_m(y)) + \mathcal{O}(\epsilon_N^{2/3}) \\ R_m^- &= \frac{-2^{4m}\nu^{2(m-1)/3}e^{-p_{m-1}}\mathcal{V}_m(y)}{2^{8(m-1)}\nu^{-2+4m/3}\epsilon_N^{2/3}e^{-2p_{m-1}}B_{m,21}(y)^2 + 16} + \mathcal{O}(\epsilon_N^{2/3}e^{-p_{m-1}}\mathcal{V}_m(y)) + \mathcal{O}(\epsilon_N^{2/3}).\end{aligned}\tag{7.41}$$

The explicit terms we have retained in  $R_m^\pm$  have the same form as  $\dot{R}_m^\pm$  defined by (7.32) and (7.35) but written in terms of the “fast” variables  $p_m$ ,  $p_{m-1}$ , and  $y$  instead of  $q_m$ ,  $q_{m-1}$ , and  $z$ . But, recalling the definition (7.29) we deduce from (4.38) that

$$p_m = q_m + \mathcal{O}(\epsilon_N^{-1}\Delta x^2, \epsilon_N^{-1}t\Delta x, \epsilon_N^{-1}t^2) = q_m + \mathcal{O}(\epsilon_N^{1/3}),\tag{7.42}$$

and from this it follows (using (7.40) and the fact that the denominators are bounded away from zero) that

$$\begin{aligned}R_m^+ &= \frac{-2^{4(1-m)}\nu^{-2m/3}e^{q_m}\mathcal{U}_m(y)}{2^{-8m}\nu^{-2(1+2m)/3}\epsilon_N^{2/3}e^{2q_m}B_{m,12}(y)^2 + 16} + \mathcal{O}(\epsilon_N^{1/3}e^{p_m}\mathcal{U}_m(y)) + \mathcal{O}(\epsilon_N^{2/3}) \\ R_m^- &= \frac{-2^{4m}\nu^{2(m-1)/3}e^{-q_{m-1}}\mathcal{V}_m(y)}{2^{8(m-1)}\nu^{-2+4m/3}\epsilon_N^{2/3}e^{-2q_{m-1}}B_{m,21}(y)^2 + 16} + \mathcal{O}(\epsilon_N^{1/3}e^{-p_{m-1}}\mathcal{V}_m(y)) + \mathcal{O}(\epsilon_N^{2/3}).\end{aligned}\tag{7.43}$$

Also, from the definition (1.43), the definition  $y = r(x, t)/\epsilon_N^{2/3}$  and (4.38), we see that

$$y = z + \mathcal{O}(\epsilon_N^{-2/3}\Delta x^2, \epsilon_N^{-2/3}t\Delta x, \epsilon_N^{-2/3}t^2) = z + \mathcal{O}(\epsilon_N^{2/3}).\tag{7.44}$$

Since the poles are simple for  $\mathcal{U}_m$  and  $\mathcal{V}_m$ ,

$$\begin{aligned}\mathcal{U}_m(y) - \mathcal{U}_m(z) &= \mathcal{O}(\epsilon_N^{2/3}\mathcal{U}_m'(y + \mathcal{O}(\epsilon_N^{2/3}))) = \mathcal{O}(\epsilon_N^{2/3}|y + \mathcal{O}(\epsilon_N^{2/3}) - \mathcal{P}(\mathcal{U}_m)|^{-2}) \\ \mathcal{V}_m(y) - \mathcal{V}_m(z) &= \mathcal{O}(\epsilon_N^{2/3}\mathcal{V}_m'(y + \mathcal{O}(\epsilon_N^{2/3}))) = \mathcal{O}(\epsilon_N^{2/3}|y + \mathcal{O}(\epsilon_N^{2/3}) - \mathcal{P}(\mathcal{V}_m)|^{-2}).\end{aligned}\tag{7.45}$$

Now,  $|y - \mathcal{P}(\mathcal{U}_m)|^{-1} = \mathcal{O}(\epsilon_N^{-1/6})$  for  $(y, s) \in \Omega_m^+ \cup \Omega_m^-$ , so

$$\begin{aligned} \frac{1}{|y + \mathcal{O}(\epsilon_N^{2/3}) - \mathcal{P}(\mathcal{U}_m)|} &= \frac{1}{|y - \mathcal{P}(\mathcal{U}_m)|} \frac{|y - \mathcal{P}(\mathcal{U}_m)|}{|y + \mathcal{O}(\epsilon_N^{2/3}) - \mathcal{P}(\mathcal{U}_m)|} \\ &\leq \frac{1}{|y - \mathcal{P}(\mathcal{U}_m)|} \frac{|y - \mathcal{P}(\mathcal{U}_m)|}{||y - \mathcal{P}(\mathcal{U}_m)| - |\mathcal{O}(\epsilon_N^{2/3})||} \\ &= \frac{1}{|y - \mathcal{P}(\mathcal{U}_m)|} \frac{1}{1 - |\mathcal{O}(\epsilon_N^{2/3})||y - \mathcal{P}(\mathcal{U}_m)|^{-1}} \\ &= \frac{1}{|y - \mathcal{P}(\mathcal{U}_m)|} \frac{1}{1 - |\mathcal{O}(\epsilon_N^{1/2})|} \end{aligned} \quad (7.46)$$

so that  $\mathcal{U}_m(y) - \mathcal{U}_m(z) = \mathcal{O}(\epsilon_N^{2/3}|y - \mathcal{P}(\mathcal{U}_m)|^{-2})$  and similarly  $\mathcal{V}_m(y) - \mathcal{V}_m(z) = \mathcal{O}(\epsilon_N^{2/3}|y - \mathcal{P}(\mathcal{V}_m)|^{-2})$ . For  $B_{m,12}^2$  and  $B_{m,21}^2$  the poles are double, and it follows by the same reasoning that  $B_{m,12}(y)^2 - B_{m,12}(z)^2 = \mathcal{O}(\epsilon_N^{2/3}|y - \mathcal{P}(\mathcal{U}_m)|^{-3})$  and  $B_{m,21}(y)^2 - B_{m,21}(z)^2 = \mathcal{O}(\epsilon_N^{2/3}|y - \mathcal{P}(\mathcal{V}_m)|^{-3})$ . From these considerations we have that

$$\begin{aligned} R_m^+ &= \dot{R}_m^+ + \mathcal{O}(\epsilon_N^{2/3} e^{p_m} |y - \mathcal{P}(\mathcal{U}_m)|^{-2}) + \mathcal{O}(\epsilon_N^{4/3} e^{3p_m} \mathcal{U}_m(y) |y - \mathcal{P}(\mathcal{U}_m)|^{-3}) \\ &\quad + \mathcal{O}(\epsilon_N^2 e^{3p_m} |y - \mathcal{P}(\mathcal{U}_m)|^{-5}) + \mathcal{O}(\epsilon_N^{1/3} e^{p_m} \mathcal{U}_m(y)) + \mathcal{O}(\epsilon_N^{2/3}), \quad (y, s) \in \Omega_m^+ \\ R_m^- &= \dot{R}_m^- + \mathcal{O}(\epsilon_N^{2/3} e^{-p_{m-1}} |y - \mathcal{P}(\mathcal{V}_m)|^{-2}) + \mathcal{O}(\epsilon_N^{4/3} e^{-3p_{m-1}} \mathcal{V}_m(y) |y - \mathcal{P}(\mathcal{V}_m)|^{-3}) \\ &\quad + \mathcal{O}(\epsilon_N^2 e^{-3p_{m-1}} |y - \mathcal{P}(\mathcal{V}_m)|^{-5}) + \mathcal{O}(\epsilon_N^{1/3} e^{-p_{m-1}} \mathcal{V}_m(y)) + \mathcal{O}(\epsilon_N^{2/3}), \quad (y, s) \in \Omega_m^-, \end{aligned} \quad (7.47)$$

where  $\dot{R}_m^+$  is defined by (7.32) and  $\dot{R}_m^-$  is defined by (7.35). By comparing with the definition (6.69) of  $e_m^+(y, s; \epsilon_N)$  for  $(y, s) \in \Omega_m^+$  and the definition (6.70) of  $e_m^-(y, s; \epsilon_N)$  for  $(y, s) \in \Omega_m^-$  we see that

$$R_m^\pm = \dot{R}_m^\pm + \mathcal{O}(e_m^\pm(y, s; \epsilon_N)), \quad (y, s) \in \Omega_m^\pm. \quad (7.48)$$

Finally, we insert the estimate (7.48) into  $\dot{C}_{N,m}^\pm(x, t)$  and  $\dot{S}_{N,m}^\pm(x, t)$  given by (7.18) noting that derivatives of the rational expressions in  $R_m^\pm$  that appear are uniformly bounded to complete the proof of the proposition.  $\square$

Combining Proposition 7.1 with the formulae (7.12), we obtain the estimates

$$\cos(\tfrac{1}{2}u_N(x, t)) = \ddot{C}_{N,m}^\pm(x, t) + \mathcal{O}(e_m^\pm(y, s; \epsilon_N)) \quad \text{and} \quad \sin(\tfrac{1}{2}u_N(x, t)) = \ddot{S}_{N,m}^\pm(x, t) + \mathcal{O}(e_m^\pm(y, s; \epsilon_N)) \quad (7.49)$$

holding uniformly for  $(y, s) \in \Omega_m^\pm$  with  $y$  bounded.

We are now in a position to define the multiscale asymptotic formulae for  $\cos(\tfrac{1}{2}u_N(x, t))$  and  $\sin(\tfrac{1}{2}u_N(x, t))$  mentioned in the statement of the Theorem:

$$\dot{C}(x, t; \epsilon_N) := \ddot{C}_{N,m}^\pm(x, t) \quad \text{and} \quad \dot{S}(x, t; \epsilon_N) := \ddot{S}_{N,m}^\pm(x, t), \quad \text{whenever } (\epsilon_N^{-2/3}r(x, t), s(x, t)) \in \Omega_m^\pm. \quad (7.50)$$

Using (7.49) we find that

$$\begin{aligned} \cos(\tfrac{1}{2}u_N(x, t)) &= \dot{C}(x, t; \epsilon_N) + \mathcal{O}(e_m^\pm(\epsilon_N^{-2/3}r(x, t), s(x, t); \epsilon_N)) = \dot{C}(x, t; \epsilon_N) + \mathcal{O}(\epsilon_N^{1/6}) \\ \sin(\tfrac{1}{2}u_N(x, t)) &= \dot{S}(x, t; \epsilon_N) + \mathcal{O}(e_m^\pm(\epsilon_N^{-2/3}r(x, t), s(x, t); \epsilon_N)) = \dot{S}(x, t; \epsilon_N) + \mathcal{O}(\epsilon_N^{1/6}) \end{aligned} \quad (7.51)$$

hold whenever  $(x, t)$  are such that  $(\epsilon_N^{-2/3}r(x, t), s(x, t)) \in \Omega_m^\pm$ . The exponent of  $1/6$  is the worst-case exponent as explained in the paragraph following the proof of Proposition 6.5. On the other hand, it is clear that this error estimate is uniformly valid whenever  $m$  is bounded and  $\epsilon_N^{-2/3}r(x, t) = y$  is bounded, or as is the same, whenever  $t = \mathcal{O}(\epsilon_N \log(\epsilon_N^{-1}))$  and  $\Delta x = \mathcal{O}(\epsilon_N^{2/3})$ . This completes the proof of Theorem 1.2.

**7.4. Proof of Theorem 1.3 (Superluminal kink asymptotics).** Choose any integer  $m$  with  $|m| \leq B$ , and let  $K > 0$  be fixed. Suppose that  $(x, t)$  lies in the rectangle  $\mathcal{R}_m$  defined by the inequalities  $|\Delta x| \leq K\epsilon_N^{2/3}$  and  $|t - \frac{2}{3}m\epsilon_N \log(\epsilon_N^{-1})| \leq \frac{1}{3}\epsilon_N \log(\epsilon_N^{-1})$ . Because  $y = z + \mathcal{O}(\epsilon_N^{2/3})$  and  $p_m = q_m + \mathcal{O}(\epsilon_N^{1/3})$ , for any *strictly positive* choice of the parameter  $\kappa \geq 0$  entering into Definition 6.1, the image  $I(\mathcal{R}_m)$  of  $\mathcal{R}_m$  in the  $(y, s)$ -plane under the mapping  $y = r(x, t)/\epsilon_N^{2/3}$  and  $s = s(x, t)$  will be contained in the horizontal strip  $\Omega_m^+ \cup \Omega_{m+1}^-$

(see (6.14)) with  $y$  bounded, as long as  $\epsilon_N$  is sufficiently small. Also, for  $(x, t)$  mapping into  $I(\mathcal{R}_m) \cap \Omega_m^+$  (respectively into  $I(\mathcal{R}_m) \cap \Omega_{m+1}^-$ ) we have the uniform estimate  $e_m^+(\epsilon_N^{-2/3} r(x, t), s(x, t); \epsilon_N) = \mathcal{O}(\epsilon_N^{1/6})$  (respectively we have the uniform estimate  $e_{m+1}^-(\epsilon_N^{-2/3} r(x, t), s(x, t); \epsilon_N) = \mathcal{O}(\epsilon_N^{1/6})$ ).

Moreover, given an interval  $[z_-, z_+]$  with  $2\nu^{1/3}|z_{\pm}| \leq K$  on which  $\log |\mathcal{U}_m(z)|$  is bounded, if the parameter  $\delta > 0$  entering into Definition 6.1 is chosen sufficiently small, the sub-rectangle  $\mathcal{R}_m^{\text{sub}} \subset \mathcal{R}_m$  defined by the inequalities  $z_- \leq z \leq z_+$  (recall the definition (1.43) of  $z$  in terms of  $\Delta x$ ) and  $|t - \frac{2}{3}m\epsilon_N \log(\epsilon_N^{-1})| \leq \frac{1}{3}\epsilon_N \log(\epsilon_N^{-1})$  will be mapped into a subregion  $I(\mathcal{R}_m^{\text{sub}})$  of  $I(\mathcal{R}_m)$  in which both inequalities  $|y - \mathcal{P}(\mathcal{U}_m)| \geq \delta$  and  $|y - \mathcal{P}(\mathcal{V}_{m+1})| \geq \delta$  hold. It then follows that for  $(x, t)$  mapping into  $I(\mathcal{R}_m^{\text{sub}}) \cap \Omega_m^+$ , the estimate  $e_m^+(\epsilon_N^{-2/3} r(x, t), s(x, t); \epsilon_N) = \mathcal{O}(\epsilon_N^{1/3})$  holds uniformly while for  $(x, t)$  mapping into  $I(\mathcal{R}_m^{\text{sub}}) \cap \Omega_{m+1}^-$ , the estimate  $e_{m+1}^-(\epsilon_N^{-2/3} r(x, t), s(x, t); \epsilon_N) = \mathcal{O}(\epsilon_N^{1/3})$  holds uniformly. See also Figure 6.5.

Suppose that  $(x, t)$  corresponds to a point  $(y, s) \in \Omega_m^+$ . We wish to neglect the first term in the denominator of  $\dot{R}_m^+$  as defined by (7.32); we write  $\dot{R}_m^+$  in the form

$$\dot{R}_m^+ = \frac{-\text{sgn}(\mathcal{U}_m(z))e^{T_K}}{f^2 + 1}, \quad f = \mathcal{O}(\epsilon_N^{1/3} e^{q_m} B_{m,12}(z)) = \mathcal{O}(\epsilon_N^{1/3} e^{q_m} |z - \mathcal{P}(\mathcal{U}_m)|^{-1}), \quad (7.52)$$

where  $T_K$  is defined in terms of  $t$ ,  $\epsilon_N$ ,  $\nu$ , and  $|\mathcal{U}_m(z)|$  by (1.42). Note that since  $z = y + \mathcal{O}(\epsilon_N^{2/3})$  and  $q_m = p_m + \mathcal{O}(\epsilon_N^{1/3})$ , and since  $|y - \mathcal{P}(\mathcal{U}_m)| \geq \delta \epsilon_N^{1/6}$  for  $(y, s) \in \Omega_m^+$ , the upper bound for  $f$  may be replaced with  $f = \mathcal{O}(\epsilon_N^{1/3} e^{p_m} |y - \mathcal{P}(\mathcal{U}_m)|^{-1})$ . Therefore, as pointed out in the proof of Proposition 7.1,  $f = \mathcal{O}(1)$  uniformly for  $(y, s) \in \Omega_m^+$ , although it is typically smaller pointwise. By using this result in (7.31) and recalling (7.49), it follows that whenever  $(x, t) \in \mathcal{R}_m$  maps into  $\Omega_m^+$ ,

$$\begin{aligned} \cos(\tfrac{1}{2}u_N(x, t)) &= (-1)^m \text{sgn}(\mathcal{U}_m(z)) \text{sech}(T_K) + \mathcal{O}(e^{-|T_K|} f^2) + \mathcal{O}(e_m^+(\epsilon_N^{-2/3} r(x, t), s(x, t); \epsilon_N)) \\ \sin(\tfrac{1}{2}u_N(x, t)) &= (-1)^{m+1} \tanh(T_K) + \mathcal{O}(e^{-2|T_K|} f^2) + \mathcal{O}(e_m^+(\epsilon_N^{-2/3} r(x, t), s(x, t); \epsilon_N)). \end{aligned} \quad (7.53)$$

As shown at the beginning of the proof, the error terms involving  $e_m^+$  are either of magnitude  $\mathcal{O}(\epsilon_N^{1/6})$  and in particular  $\mathfrak{o}(1)$ , or, if  $(x, t) \in \mathcal{R}_m^{\text{sub}}$ , of smaller magnitude  $\mathcal{O}(\epsilon_N^{1/3})$ . The remaining error terms are bounded, and they are negligible unless *both*  $T_K = \mathcal{O}(1)$  and  $f$  fails to be small. But since  $f = \mathcal{O}(\epsilon_N^{1/3} e^{p_m} |y - \mathcal{P}(\mathcal{U}_m)|^{-1})$  and  $(y, s) \in \Omega_m^+$ , we see from Definition 6.1 that  $f$  fails to be small only if  $e^{-p_m} = \mathcal{O}(\epsilon_N^{1/3})$ , or equivalently,  $e^{-q_m} = \epsilon_N^{-2m/3} e^{-t/\epsilon_N} = \mathcal{O}(\epsilon_N^{1/3})$ . Now, under this condition we have from (1.42) that  $T_K \geq \mathcal{O}(1) + \frac{1}{3} \log(\epsilon_N^{-1}) + \log |\mathcal{U}_m(z)|$ , so to have  $T_K = \mathcal{O}(1)$  requires that  $|z - \mathcal{P}(\mathcal{U}_m)| = \mathcal{O}(\epsilon_N^{1/3})$ . Moreover, if  $(x, t) \in \mathcal{R}_m^{\text{sub}}$  and its image  $(y, s) = (\epsilon_N^{-2/3} r(x, t), s(x, t))$  lies in  $\Omega_m^+$ ,  $e^{-|T_K|} = \mathcal{O}(e^{-|q_m|})$  and  $f = \mathcal{O}(\epsilon_N^{1/3} e^{q_m})$ . The inequality  $|q_m| \leq \frac{1}{3} \log(\epsilon_N^{-1})$  then implies that if  $(x, t) \in \mathcal{R}_m^{\text{sub}}$  with image in  $\Omega_m^+$ , then  $e^{-|T_K|} f^2$  and  $e^{-2|T_K|} f^2$  are both dominated by the  $\mathcal{O}(\epsilon_N^{1/3})$  estimate of the error terms involving  $e_m^+$ .

Now suppose instead that  $(x, t)$  corresponds to a point  $(y, s) \in \Omega_{m+1}^-$ . As before, we wish to neglect the first term in the denominator of  $\dot{R}_{m+1}^-$  as defined by (7.35), so we write  $\dot{R}_{m+1}^-$  in the form (using the identity  $\mathcal{V}_{m+1}(z) = \mathcal{U}_m(z)^{-1}$ )

$$\dot{R}_{m+1}^- = \frac{-\text{sgn}(\mathcal{U}_m(z))e^{-T_K}}{h^2 + 1}, \quad h = \mathcal{O}(\epsilon_N^{1/3} e^{-q_m} B_{m+1,21}(z)) = \mathcal{O}(\epsilon_N^{1/3} e^{-q_m} |z - \mathcal{P}(\mathcal{U}_m)|^{-1}). \quad (7.54)$$

Again, the error bound for  $h$  may be replaced with  $\mathcal{O}(\epsilon_N^{1/3} e^{-p_m} |y - \mathcal{P}(\mathcal{U}_m)|^{-1})$  and so  $h = \mathcal{O}(1)$  holds uniformly for  $(y, s) \in \Omega_{m+1}^-$ ; it then follows by using this result in (7.34) and recalling (7.49), that whenever  $(x, t) \in \mathcal{R}_m$  maps into  $\Omega_{m+1}^-$ ,

$$\begin{aligned} \cos(\tfrac{1}{2}u_N(x, t)) &= (-1)^m \text{sgn}(\mathcal{U}_m(z)) \text{sech}(T_K) + \mathcal{O}(e^{-|T_K|} h^2) + \mathcal{O}(e_{m+1}^-(\epsilon_N^{-2/3} r(x, t), s(x, t); \epsilon_N)) \\ \sin(\tfrac{1}{2}u_N(x, t)) &= (-1)^{m+1} \tanh(T_K) + \mathcal{O}(e^{-2|T_K|} h^2) + \mathcal{O}(e_{m+1}^-(\epsilon_N^{-2/3} r(x, t), s(x, t); \epsilon_N)). \end{aligned} \quad (7.55)$$

In particular, we note that the leading terms are *identical* with those arising from the analysis in the case that the image of  $(x, t)$  lies in  $\Omega_m^+$ . By completely analogous reasoning as in that case, one learns that the sum of error terms is  $\mathfrak{o}(1)$  unless both  $T_K = \mathcal{O}(1)$  and also  $h$  fails to be small, which occurs in this case only if  $e^{p_m} = \mathcal{O}(\epsilon_N^{1/3})$ , or equivalently, if  $e^{q_m} = \epsilon_N^{2m/3} e^{t/\epsilon_N} = \mathcal{O}(\epsilon_N^{1/3})$ . Under this condition, the definition (1.42)

implies that  $T_K \leq \mathcal{O}(1) - \frac{1}{3} \log(\epsilon_N^{-1}) + \log |\mathcal{U}_m(z)|$ , so to have  $T_K = \mathcal{O}(1)$  requires that  $|z - \mathcal{P}(\mathcal{U}_m)| = \mathcal{O}(\epsilon_N^{1/3})$ . Moreover, if  $(x, t) \in \mathcal{R}_m^{\text{sub}}$  with image in  $\Omega_{m+1}^-$ , then both error terms are  $\mathcal{O}(\epsilon_N^{1/3})$ . This completes the proof of Theorem 1.3.

**7.5. Proof of Theorem 1.4 (Grazing kink collisional asymptotics).** Recall that  $z_0 \in \mathcal{Z}(\mathcal{U}_{m-1})$  is a simple zero of  $\mathcal{U}_{m-1}$ , and hence also a simple zero of  $\mathcal{V}_{m+1}$ . Note that from the recurrence relation (5.62) it follows that  $\mathcal{V}'_{m+1}(z_0) = -1/\mathcal{U}'_{m-1}(z_0)$ . Since  $z = y + \mathcal{O}(\epsilon_N^{2/3})$  and  $t = s + \mathcal{O}(\epsilon_N^{4/3})$  for the range of  $x$  and  $t$  under consideration, the region in the  $(x, t)$ -plane defined by the inequalities

$$|t - (\frac{2}{3}m - \frac{1}{3})\epsilon_N \log(\epsilon_N^{-1})| \leq \frac{1}{3}\epsilon_N \log(\epsilon_N^{-1}) \quad \text{and} \quad |z - z_0| \leq \mu \epsilon_N^{1/6} \exp(|t - (\frac{2}{3}m - \frac{1}{3})\epsilon_N \log(\epsilon_N^{-1})|/(2\epsilon_N)) \quad (7.56)$$

is, for both  $\mu > 0$  and  $\epsilon_N$  sufficiently small, mapped into the union  $\Omega_{m-1}^+ \cup \Omega_{m+1}^-$  in the  $(y, s)$ -plane, where  $(y, s) = (\epsilon_N^{-2/3}r(x, t), s(x, t))$ . In particular, each point of the image lies either in the upward-pointing “tooth” of  $\Omega_{m-1}^+$  centered at  $y = z_0$  or in the downward-pointing “tooth” of  $\Omega_{m+1}^-$  centered at  $y = z_0$ . It follows that for the points whose images lie in  $\Omega_{m-1}^+$ , the error term  $e_{m-1}^+(\epsilon_N^{-2/3}r(x, t), s(x, t); \epsilon)$  is  $\mathcal{O}(\epsilon_N^{1/6})$  while for the points whose images lie in  $\Omega_{m+1}^-$ , the error term  $e_{m+1}^-(\epsilon_N^{-2/3}r(x, t), s(x, t); \epsilon_N)$  is also  $\mathcal{O}(\epsilon_N^{1/6})$  so that in particular these error terms arising in (7.49) will always tend to zero with  $\epsilon_N$ . Moreover, the inequality  $|z - z_0| = \mathcal{O}(\epsilon_N^{1/3})$  implies that  $|y - z_0| = \mathcal{O}(\epsilon_N^{1/3})$ , from which it follows that under this condition  $(x, t)$  will be mapped into the parts of  $\Omega_{m-1}^+$  and  $\Omega_{m+1}^-$  (the centers of two matching teeth) in which the better error estimate of  $\mathcal{O}(\epsilon_N^{1/3})$  holds for  $e_{m-1}^+$  and  $e_{m+1}^-$  respectively.

Suppose that the image of  $(x, t)$  in the  $(y, s)$ -plane lies in the upward-pointing tooth of  $\Omega_{m-1}^+$  centered at  $y = z_0$ . In this part of the image, the uniform estimate  $e^{T_G} = \mathcal{O}(1)$  holds. By Taylor expansion,

$$\begin{aligned} \mathcal{U}_{m-1}(z) &= \mathcal{U}'_{m-1}(z_0)(z - z_0)(1 + \mathcal{O}(z - z_0)) = \frac{\epsilon_N^{1/3}\mathcal{U}'_{m-1}(z_0)}{2\nu^{1/3}}X_G(1 + \mathcal{O}(z - z_0)) \\ B_{m-1,12}(z) &= 2H_{m-1}(z)\mathcal{U}_{m-1}(z) - \mathcal{U}'_{m-1}(z) = -\mathcal{U}'_{m-1}(z_0) + \mathcal{O}(z - z_0), \end{aligned} \quad (7.57)$$

where the spatial coordinate  $X_G$  is defined by (1.49). It follows from (7.32) using  $e^{T_G} = \mathcal{O}(1)$  that

$$\dot{R}_{m-1}^+ = -\text{sgn}(\mathcal{U}'_{m-1}(z_0))X_G \text{sech}(T_G)(1 + \mathcal{O}(z - z_0)), \quad (7.58)$$

and then that

$$\begin{aligned} \ddot{C}_{N,m-1}^+(x, t) &= (-1)^{m-1} \text{sgn}(\mathcal{U}'_{m-1}(z_0)) \frac{2X_G \text{sech}(T_G)}{1 + X_G^2 \text{sech}^2(T_G)} + \mathcal{O}(z - z_0) \\ \ddot{S}_{N,m-1}^+(x, t) &= (-1)^{m-1} \frac{1 - X_G^2 \text{sech}^2(T_G)}{1 + X_G^2 \text{sech}^2(T_G)} + \mathcal{O}(z - z_0). \end{aligned} \quad (7.59)$$

Combining this formula with the above estimates of  $e_{m-1}^+(x, t; \epsilon_N)$  and the formulae (7.49) proves the Theorem for those  $(x, t)$  whose image in the  $(y, s)$ -plane lies within the upward-pointing tooth of  $\Omega_{m-1}^+$  centered at  $y = z_0$ .

Now suppose instead that the image of  $(x, t)$  in the  $(y, s)$ -plane lies in the downward-pointing tooth of  $\Omega_{m+1}^-$  centered at  $y = z_0$ . In this part of the image, the uniform estimate  $e^{-T_G} = \mathcal{O}(1)$  holds. By Taylor expansion and the fact that  $\mathcal{V}'_{m+1}(z_0) = -1/\mathcal{U}'_{m-1}(z_0)$ ,

$$\begin{aligned} \mathcal{V}_{m+1}(z) &= -\frac{1}{\mathcal{U}'_{m-1}(z_0)}(z - z_0)(1 + \mathcal{O}(z - z_0)) = -\frac{\epsilon_N^{1/3}}{2\nu^{1/3}\mathcal{U}'_{m-1}(z_0)}X_G(1 + \mathcal{O}(z - z_0)) \\ B_{m+1,21}(z) &= \mathcal{V}'_{m+1}(z) - 2H_{m+1}(z)\mathcal{V}_{m+1}(z) = -\frac{1}{\mathcal{U}'_{m-1}(z_0)} + \mathcal{O}(z - z_0). \end{aligned} \quad (7.60)$$

It then follows from (7.35) using  $e^{-T_G} = \mathcal{O}(1)$  that

$$\dot{R}_{m+1}^- = \text{sgn}(\mathcal{U}'_{m-1}(z_0))X_G \text{sech}(T_G)(1 + \mathcal{O}(z - z_0)), \quad (7.61)$$

and then that

$$\begin{aligned}\ddot{e}_{N,m+1}^-(x,t) &= (-1)^{m-1} \operatorname{sgn}(U'_{m-1}(z_0)) \frac{2X_G \operatorname{sech}(T_G)}{1 + X_G^2 \operatorname{sech}^2(T_G)} + \mathcal{O}(z - z_0) \\ \ddot{S}_{N,m+1}^-(x,t) &= (-1)^{m-1} \frac{1 - X_G^2 \operatorname{sech}^2(T_G)}{1 + X_G^2 \operatorname{sech}^2(T_G)} + \mathcal{O}(z - z_0).\end{aligned}\tag{7.62}$$

Combining this formula with the above estimates of  $e_{m+1}^-(x, t; \epsilon_N)$  and the formulae (7.49) proves the Theorem for those  $(x, t)$  whose image in the  $(y, s)$ -plane lies within the downward-pointing tooth of  $\Omega_{m+1}^-$  centered at  $y = z_0$ .

## REFERENCES

- [1] M. Ablowitz, D. Kaup, A. Newell, and H. Segur, “Method for solving the sine-Gordon equation,” *Phys. Rev. Lett.*, **30**, 1262–1264, 1973.
- [2] J. Baik, P. Deift, and K. Johansson, “On the distribution of the length of the longest increasing subsequence of random permutations,” *J. AMS*, **12**, 1119–1178, 1999.
- [3] A. Barone, F. Esposito, C. J. Magee, and A. C. Scott, “Theory and applications of the sine-Gordon equation,” *Rev. Nuovo Cim.*, **1**, 227–267, 1971.
- [4] M. Bertola and A. Tovbis, “Universality in the profile of the semiclassical limit solutions to the focusing nonlinear Schrödinger equation at the first breaking curve,” *Int. Math. Res. Not.*, **2010**, 2119–2167, 2010.
- [5] M. Bertola and A. Tovbis, “Universality for the focusing nonlinear Schrödinger equation at the gradient catastrophe point: rational breathers and poles of the tritronquée solution to Painlevé I,” [arXiv:1004.1828](#), 2010.
- [6] R. J. Buckingham and P. D. Miller, “Exact solutions of semiclassical non-characteristic Cauchy problems for the sine-Gordon equation,” *Physica D*, **237**, 2296–2341, 2008.
- [7] R. J. Buckingham and P. D. Miller, “The sine-Gordon equation in the semiclassical limit: dynamics of fluxon condensates,” [arXiv:1103.0061](#), 2011.
- [8] P. Cheng, S. Venakides, and X. Zhou, “Long-time asymptotics for the pure radiation solution of the sine-Gordon equation,” *Comm. Part. Diff. Eq.*, **24**, 1195–1262, 1999.
- [9] C. Chester, B. Friedman, and F. Ursell, “An extension of the method of steepest descents,” *Proc. Cambridge Philos. Soc.*, **53**, 599–611, 1957.
- [10] T. Claeys and T. Grava, “Painlevé II asymptotics near the leading edge of the oscillatory zone for the Korteweg-de Vries equation in the small-dispersion limit,” *Comm. Pure Appl. Math.*, **63**, 203–232, 2010.
- [11] T. Claeys and T. Grava, “Solitonic asymptotics for the Korteweg-de Vries equation in the small dispersion limit,” *SIAM J. Math. Anal.*, **42**, 2132–2154, 2010.
- [12] T. Claeys and A. Kuijlaars, “Universality of the double scaling limit in random matrix models,” *Comm. Pure Appl. Math.*, **59**, 1573–1603, 2006.
- [13] P. A. Clarkson, “Remarks on the Yablonskii-Vorob’ev polynomials,” *Phys. Lett. A*, **319**, 137–144, 2003.
- [14] P. A. Clarkson and E. L. Mansfield, “The second Painlevé equation, its hierarchy and associated special polynomials,” *Nonlinearity*, **16**, R1–R26, 2003.
- [15] P. Deift, S. Venakides, and X. Zhou, “New results in small dispersion KdV by an extension of the steepest descent method for Riemann-Hilbert problems,” *Internat. Math. Res. Notices*, **1997**, 285–299, 1997.
- [16] P. Deift and X. Zhou, “A steepest-descent method for oscillatory Riemann-Hilbert problems: asymptotics for the mKdV equation,” *Ann. of Math.*, **137**, 295–368, 1993.
- [17] L. Faddeev and L. Takhtajan, *Hamiltonian Methods in the Theory of Solitons*, Springer-Verlag, New York, 1987.
- [18] A. S. Fokas, A. R. Its, A. A. Kapaev, and V. Yu. Novokshenov, *Painlevé Transcendents: The Riemann-Hilbert Approach*, Volume 128, Mathematical Surveys and Monographs, American Mathematical Society, Providence, 2006.
- [19] E. L. Ince, *Ordinary Differential Equations*, Dover Publications, New York, 1956.
- [20] D. Kaup, “Method for solving the sine-Gordon equation in laboratory coordinates,” *Stud. Appl. Math.*, **54**, 165–179, 1975.
- [21] M. Klaus and J. K. Shaw, “Purely imaginary eigenvalues of Zakharov-Shabat systems,” *Phys. Rev. E*, **65**, 36607–36611, 2002.
- [22] Y. Murata, “Rational solutions of the second and fourth Painlevé equations,” *Funkcialaj Ekvacioj*, **28**, 1–32, 1985.
- [23] A. C. Scott, F. Chu, S. Reible, “Magnetic flux propagation on a Josephson transmission line,” *J. Appl. Phys.*, **47**, 3272–3286, 1976.
- [24] V. E. Zakharov and A. B. Shabat, “Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media,” *Sov. Phys. JETP*, **34**, 62–69, 1972.
- [25] V. E. Zakharov, L. Takhtajan, and L. Faddeev, “Complete description of solutions of the sine-Gordon equation,” *Dokl. Akad. Nauk. SSSR*, **219**, 1334–1337, 1974.
- [26] X. Zhou, “Inverse scattering transform for systems with rational spectral dependence,” *J. Diff. Eq.*, **115**, 277–303, 1995.